Session Papers for the 21st Annual National Conference
on Managing Environmental Quality Systems
April 8-11, 2002
Phoenix, Arizona

APPLIED QUALITY ASSURANCE IN RESEARCH

Contributions of Quality Assurance/Quality Control to Emerging Techniques in Ecology: Not Just For Chemists Anymore - N. Detenbeck, U.S. EPA

Development of an ETD Surveillance Checklist For Monitoring EPA Research Activities - T. Hughes, U.S. EPA

The Importance of a Successful Quality Assurance Program From a Research Manager's Perspective - W. Ponder, U.S. EPA

Quality Assurance Project Plans: A Useless Paper Exercise or Valuable Aid? - B. Schumacher, U.S. EPA

Successful Quality Assurance (QA) Programs in EPA's Office of Research and Development Laboratories - B. Culpepper, U.S. EPA

DEFINING PERFORMANCE CRITERIA

Establishing Sensitivity Requirements for Environmental Analyses for Project Data Quality Objectives - T. Georgian, C.R. Mao, U.S. Army Corps of Engineers

Getting to the Bottom Line: Decision Quality vs. Data Quality - D. Crumbling, U.S. EPA

Performance Based Approach and Data Quality - B. Hull and J. Griggs, U.S. EPA

Quantifying Uncertainty: Are We There Yet? - L. Blume, U.S. EPA; J. Schofield and K. Miller, DynCorp I&ET, Inc.

The Role of Quality Assurance in the Evaluation of Two Landfill Bioreactor Operational Techniques at an Existing Landfill - J.T. Markwiese, A.M. Vega, R. Green, P. Black

Working on a Quality Improvement Case Study: Sampling and Analysis Plans Under the Navy Installation Restoration Program - N.A. Ancog, Naval Facilities Engineering Command; J. Dirgo, R. Ohta, G. Swanson, Tetratech

DOCUMENTS AND RECORDS MANAGEMENT

Adventures in Environmental Data Reporting: High Tech, Low Tech and Everything In Between - D. Dinsmore, Wisconsin DNR

The NEIC Document Control System - K. Mathews, U.S. EPA

ENVIRONMENTAL LABORATORY DATA AND INFORMATION

Accreditation at the U.S. EPA - NEIC - B. Hughes, K.E. Nottingham, J. Suggs, U.S. EPA

Improved Quality Data Systems through the Use of Standard Electronic Data - P. Wehrman, U.S. Army Corps of Engineers; R. Amano, Laboratory Data Consultants

Required Steps for the Validation of a Laboratory Information Management System - E. Turner, J. Bolton, U.S. Army Corps of Engineers

ENVIRONMENTAL QUALITY SYSTEMS

Comparison of EPA's QMS to SEI's CMMI - P. Mills, L. Braun, D. Marohl, DynCorp

A Menu of Quality Systems: From TV Dinners to Filet Mignon - W. Telliard, M. Kelly, L. Blume, U.S. EPA; H. McCarty, J. Schofield, J. Cuddeback, DynCorp Systems and Solutions

Transforming an EPA QA/R-2 Quality Management Plan into an ISO 9002 Quality Management System - R. Kell, IT Corp; E. Reynolds, U.S. EPA; C. Hedin, G. Kassakhian, IT Corporation

INFORMATION QUALITY WORK AND INITIATIVES

Enforcement and Compliance Data Quality Strategy - D. Sprague, U.S. EPA

EPA Information Quality Guidelines - E. Cummings, U.S. EPA

Understanding Enterprise Information Architectures for Managing Quality - M. Doehnert, U.S. EPA

QUALITY ASSURANCE APPLICATIONS AND EXAMPLES

Atlanta Supersite Quality Assurance Final Assessment Report - D. Mikel, U.S.EPA

Speciation Trend Network - D. Mikel, U.S. EPA

QUALITY ASSURANCE APPLICATIONS FOR AQUEOUS AND BIOLOGICAL TESTING

EPA Method 1631 Quality Control for Sampling: Closing the Loop - W. Telliard, U.S.EPA; H. McCarty, J. Schofield, DynCorp

Improving Data Confidence for EPA's Whole Effluent Toxicity Tests - C. Craig, M. Carter, S. Kassner, J. Lowry, Environmental Resource Associates

One Fish, Two Fish, We QC Fish: Controlling Data Quality Among More than 50 Organizations Over a Four Year Period - L. Riddick, DynCorp; C. Simbanin, U.S. EPA

Quality Control in Whole Effluent Toxicity Test Methods: Lessons from EPA's Interlaboratory Variability Study - W. Telliard, M. Kelly, U.S. EPA; R. Brent, H. McCarty, DynCorp

QUALITY ASSURANCE PRACTICES IN AIR MONITORING

Analysis of the OAQPS Policy on the Secondary Use of Data - R. Wright, K. Morgan, I. Beaty, RTI International

Annual Recertification Program for Audit Standards in the EPA PM2.5 Performance Evaluation Program - R. Wright, J.S. Nichol, RTI; M. Papp, P. Groff, U.S. EPA; M. Tufts, ARCADIS Geraughty & Miller

Data and Metadata Reporting Standards for the EPA's PM Supersites Research Program - L. Hook, S. Christensen, Oak Ridge National Laboratory; W. Sukloff, Environment Canada

QUALITY SYSTEM ASSESSMENTS AND PERFORMANCE MEASURES

Building the Airplane in Flight: An Auditing Approach to Quality Management System Development - M. Burson, Maine Department of Environmental Protection

Quality Performance Evaluation: Measuring Quality Management Systems Performance - G. Johnson, U.S. EPA

WORKSHOPS

The Appropriate Use of Professional Judgment in Sampling Design - M. Bertoni, RTI; J. Warren, U.S. EPA; K. Morgan, RTI

Graded Approach for Assistance Agreements Workshop

Graded Approach Workshop for Assistance Agreements
L. Blume, U.S. EPA

A Graded Approach to Documenting the Use of Existing Data by Assistance Agreement Holders - Patricia Lafornara, EPA

Improving the Quality System Specifications for EPA's Contracts - B. Young, A. Batterman, M. Doehnert, H. Ferguson, N. Parry, A. Vega, U.S. EPA

Measurement Uncertainty Expression of Sampling and Testing Data - M. Moore, Advanced Systems, Inc.

PM2.5 Ambient Air Monitoring Program: Use of Data Quality Objective Software Tool for Data Quality Assessment - M. Papp, U.S. EPA

Preparation of an SOP for an Emergency Quality Assurance Project Plan - C. Byrne, U.S. EPA

Sequential Sampling Approaches Within Visual Sample Plan (VSP) In Support of Dynamic Field Activities - B. Pulsipher, J. Wilson, D. Gilbert, N. Hassig, Pacific Northwest National Labs

Contributions of Quality Assurance/Quality Control to Emerging Techniques in Ecology: Not Just for Chemists

Naomi E. Detenbeck, EPA,. NHEER, Mid-Continent Ecology Division

The structure of biological monitoring designs has become critical as support not only for assessments of condition under Section 305(b) of the Clean Water Act but also as the starting point for site-specific determinations of impairment, diagnosis of causes of impairment, allocation of causality among pollutants and other stressors, and eventually assessment of the success of potentially costly watershed remediation plans. Quality assurance plans have historically been designed to support monitoring and analysis of chemical constituents, and the expansion of biological and habitat monitoring programs, watershed-scale monitoring designs, multivariate analysis, and landscape assessments with application of geographic information systems (GIS) have stretched traditional concepts of data quality objectives, accuracy, precision, completeness, and comparability beyond their original applications. Researchers at the Mid-Continent Ecology Division (MED-Duluth) were required to confront these issues in development of a QA plan for a three year, multi-investigator, multi-disciplinary comparative watershed project. Data quality objectives were used to determine the level of replication of watershed units and the overall experimental design, based on targets for Type I and Type II errors in establishing differences among watershed classes. A team-based approach was applied in development of both the workplan and quality assurance plan, which included descriptions of both the overall study design and subproject components. Team members identified approaches to apply QA concepts to habitat sampling, biological sampling, GIS analysis, and multivariate statistical analyses, as well as traditional chemical analyses. Lessons learned included the need for flexibility and back-up plans in developing field sampling designs and analytical approaches at the watershed scale due to complications of weather variability, site access constraints, and confounding factors relative to watershed attributes used to establish sample design strata. A standardized series of QA/QC programs were used to automate summaries of QA/QC information, but it was found that a more user-friendly interface for team chemists was needed to facilitate timely communication of QC trends and correction of problems. As new multivariate and nonparametric statistical techniques are explored, unique approaches to assess the power of designs have been identified. Finally, as data are being analyzed and interpreted, team members must grapple with the challenge of how to present estimates of uncertainty to the general public when informing them of results with policy implications for land-use decisions. In the future, early input from stakeholders will be required in the development and assessment of monitoring designs, as there will be tradeoffs among multiple objectives and assessment questions associated with streamlining of the 305(b) and 303(d) programs

Development of an ETD Surveillance Checklist for Monitoring EPA Research Activities

Thomas J. Hughes, NHEERL, EPA, Experimental Toxicology Division

Research studies conducted within the nine divisions of NHEERL are ranked in QA Categories from 1-4 based on their importance to the Agency's goals and impact on regulations. QA Category 1 studies have direct regulatory impact or very high visibility (e.g., World Trade Center research); QA Category 2 studies have potential regulatory impact or high visibility (e.g., research on black mold which has caused death in children); QA Category 3 research investigates research concepts and principles (e.g., will these assays produce a dose response after exposure to toxicants in water?); and QA Category 4 research is exploratory in nature (which bioassays will detect carcinogens in air?). The Principal Investigator, Branch Chief and QA Manager determine the category of the research when the intramural research protocol (IRP, which is equivalent to a QAPP) is written at the start of the study. The difference between a QA 1 or 2 study, or a QA 3 or 4 study can be subtle, and studies can be upgraded if the results of the research warrant. QA 1 and 2 studies are required to be audited by a QA review, called a technical systems review (TSR), during the life of the study (hopefully in the first third of the study to correct deficiencies). A TSR usually takes the division QA Manager (the lead reviewer), another scientist or QA Manager, and the laboratory staff several days to complete. A TSR for the QA Manager, from initiation of the agenda to the delivery of the final report, can take several weeks of time to complete. Priority for reviews of research is given to QA 1 and 2 studies, which can leave QA 3 and 4 studies with limited review by the QA Manager, especially in a large division such as ETD with 135 scientists and over 40 studies underway every year. It is virtually impossible to conduct a TSR on the majority of studies; reviews of ten percent of research studies in any one year is the goal. Consequently, if TSRs are the only QA review process, many research projects may not be reviewed during the life of the study. To overcome this obvious obstacle to timely QA review of research projects, an ETD surveillance checklist was developed and utilized to evaluate and review major components of all research studies within ETD. The ETD Surveillance Checklist is a condensation of the 20-page TSR checklist, is in a yes/no format, and covers notebooks, OPs, IRPs, computer files and data, data storage and filing, primary balance, primary pH meter, and two major pieces of equipment. The three page ETD Surveillance Checklist concludes with a section on exemplary findings and areas for improvement. The ETD Checklist allowed the QA Manager to quickly (one hour) and efficiently evaluate the QA status of studies for each PI in the Division. Although obviously not as thorough as a TSR, it does provide a documented basis to identify and correct deficiencies of all research studies within the Division regardless of QA Category. It is the intent of the ETD QA Manager to use this surveillance checklist on a yearly basis. This is an abstract for presentation which has been reviewed by the U.S. EPA; views expressed do not necessarily represent EPA policy.

THE IMPORTANCE OF A SUCCESSFUL QUALITY ASSURANCE (QA) PROGRAM FROM A RESEARCH MANAGER'S PERSPECTIVE

Wade H. Ponder
Chief, Technical Services Branch
EPA/ORD/NRMRL/APPCD

Research Managers' Responsibilities for QA – One responsibility of research managers is to ensure that data from research projects are acquired, processed, reported, and used in accordance with the Quality Assurance (QA) requirements established by the organization. When management does not take QA seriously or when QA requirements are not fully implemented in an organization, the results can be embarrassing, damaging, or dangerous to the organization as well as its customers or clients. The literature and news media report often the results that occur when QA is not a priority or when QA requirements are not implemented properly.

Notable QA Failures – QA has uncounted successes, but our nature is to focus on its failures. Unfortunately, there is a long list of failures to consider. Three examples have been taken from this long list to reinforce the points I wish to make about QA from a research manager's perspective:

1. On December 3, 1999, Tamara Lytle, a staff writer for "The Orlando Sentinel," published an article which stated that an Air Force report concluded that five rocket launches, from August 1998 through May 1999, failed due to cutbacks in QA staff by the contractors, poor engineering and workmanship by the contractors, and lax monitoring by U.S. military managers. Further, it was reported that one contractor had reduced QA staff on the projects by more than 60% in the years preceding the failures. The five failures cost U.S. taxpayers more than $3 billion and delayed the deployment of military payloads that were designed to help the U.S. catch terrorists as well as commercial payloads intended to enhance mobile communications capabilities.

2. The news media have provided extensive coverage of the conflicts between Ford Motor Company and the Firestone Company related to the failure of Firestone tires on Ford's Explorer Sport Utility Vehicle. Robert Polz wrote an article on tire failures in "Reliability Engineering" in which he critiqued and commented on an article entitled "Tire Failures, SUV Rollovers Put Quality on Trial," published in the December 2000 issue of American Society for Quality's ***Quality Progress*** magazine. Polz states that, if the failures are design-related, it is the province of reliability engineering to determine the root cause(s) of the failures. However, if the failures are production-related, then QA should spearhead the assessment. Since the cause is uncertain in the Ford/Firestone case, Polz concludes that both disciplines should work together as a team to determine the cause(s). Meanwhile, the National Highway Transportation Safety Administration (NHTSA) has already received more than 1400 complaints, with reports of 88 fatalities and 250 injuries.

3. On January 26, 1986, the Space Shuttle Challenger exploded during take-off. The first cousin of Mike Smith, the pilot of Challenger, at that time was a member of the NRMRL/APPCD staff. Today, more than 16 years later, I still have displayed in my office a decal commemorating the Challenger's crew and its mission. Investigations have identified the failure of o-rings in the external fuel tanks as the

cause of the catastrophic explosion that doomed the Challenger and its crew. In the hours before launch, engineers familiar with the design of the external fuel tanks and their o-rings recommended and strongly urged that the Challenger not be launched on January 26, because the ambient temperature was less than the safe minimum temperature for the o-rings to function safely and effectively. Since the engineers took these actions in an attempt to delay the launch, it may be concluded that the QA requirements for safe launch and operation of the Challenger were known but not adhered to. The result was the loss of all of Challenger's crew members, and America's Space Shuttle program was put on hold for about 2 years.

The EPA QA Analogy – So, what analogy, if any, is there between these high visibility examples in which QA was not a priority or QA requirements were not followed and QA in EPA? To be sure, EPA does not design or build rockets that might explode on their launch pads, design or build consumer products that might injure or kill users, or provide QA for products, systems, or devices, such as those on the Space Shuttle, that could fail and injure or kill people. However, it could be argued that EPA does, in fact, have analogous QA responsibilities. The Agency is charged with the protection of public health and the environment, and meeting that responsibility requires extensive research programs which generate data. It is estimated that EPA and the regulated community spend about $5 billion annually collecting environmental data. In addition to data, EPA generates information, software, and other tools which are ultimately used to justify, establish, and defend national and state standards for pollutant emissions and exposures. These emission and exposure limits are designed, first and foremost, to protect the health of the American public. If EPA had undertaken this mission and tried to see it through without adequate management and staff attention to QA, it is possible that regulatory decisions could have been made on the basis of flawed data and information and that public health and the environment might not be protected as well as they could have been.

Suggested Requirements for EPA Managers – To its credit, EPA has spent a lot of time and staff effort developing and putting in place an impressive Quality Management Plan (QMP) guidance document which delineates QA requirements for all Agency organizations involved in the acquisition, processing, and publication of research data and information. The guidance document is used by EPA's National Laboratories and Offices to prepare their QMPs. However, as indicated in the QA failure examples above, the effectiveness of any QMP is dependent on the commitment that management and research staff alike make to its implementation. From a management perspective, this commitment requires that research managers understand Agency QA requirements; establish an effective QA program to ensure that data are of known quality which is acceptable for the intended use of the data; and provide support, guidance, and oversight to principal investigators (PIs) in meeting QA requirements. It is also important that managers lead by example in the QA area:

1. Managers should make sure that other managers and PIs are aware that they view QA as an essential, integrated component of the research programs;
2. Managers should provide adequate resources (people and money) to support an effective QA program;
3. Managers should encourage collaborative, non-confrontational interactions between PIs and QA professionals; and

4. Managers should maintain oversight so that issues which have the potential for adversely affecting research and QA objectives can be negotiated and corrected quickly.

How QA Works in NRMRL/APPCD – The Air Pollution Prevention and Control Division has established a QA Team to support its four research branches. The QA Team is housed in the Technical Services Branch and consists of a team leader (who is a QA professional) and five team members (3 of whom are QA professionals). The other two members are a professional who manages the QA contract for the team and a Senior Environmental Employee who assists the Team with QA data base management and reports preparation. The four QA professionals conduct all QA reviews and audits (with contractor assistance, as needed). QA Team members share QA work among themselves to maintain a reasonable balance of workloads. They also spend up to 30% of their time working directly with the research branches' PIs helping to plan, conduct, and oversee research projects. (Of course, no QA Team member is ever allowed to conduct QA reviews of research work in which he/she has been involved.) In FY 2001, the QA Team reviewed 63 QA planning documents/research products with an average turnaround time of 5.7 days. The Team members also reviewed 121 journal articles and reports in FY 2001. To date in FY 2002, the QA Team has completed the following reviews:

	Items Reviewed in FY 2002	**Number Reviewed Through 3/19/02**
1.	Journal articles and reports	32
2.	Test plans, QAPPs, Reports, Protocols	52
3.	Funding packages	20
4.	Responses to PIs' revisions	10
5.	SOPs	2
6.	QMPs	1

When QA Team members are involved to this extent with the PIs and their research activities, there is a high probability that differences of opinion will occur. Given that, I would like to emphasize the need for managers to lead by example, specifically item 3 under **Suggested Requirements for EPA Managers,** above: i.e., "Managers should encourage cooperative, productive interactions between PIs and QA professionals." The QA philosophy in NRMRL/APPCD is that the QA Team will aid the researchers in any way possible to produce a timely, high quality product. To implement this philosophy and avoid the counterproductive trap of being viewed as the "QA police," we have borrowed and employed the five "Basic Principles" from the Zenger-Miller team training course work:

1. Focus on the situation, issue, or behavior, not on the person.
2. Maintain the self-confidence and self-esteem of others.
3. Maintain constructive relationships.
4. Take initiative to make things better.
5. Lead by example.

The use of these principles by the QA Team members in their interaction with PIs has resulted in the QA Team's maintaining the respect and cooperation of the PIs so that almost all interactions between them are positive and helpful. Infrequently, when QA Team members and PIs reach an impasse, the QA Team's Branch Chief and the PI's Branch Chief meet with the QA Team member and the PI involved in the dispute. In these meetings, it is essential that the same "Basic Principles" be employed. Using this philosophy and the "Basic Principles," only three or four meetings involving the Branch Chiefs have occurred in the last 7 years. During this time, no QA issue between a QA Team member and a PI has been elevated to the Division Director for resolution.

Conclusion – For the last 7 years, one of my responsibilities has been to supervise NRMRL/APPCD's QA Team. Through that experience, I have concluded that the collaborative, non-confrontational spirit exhibited by the members of the QA Team in their interactions with PIs and managers is a major component of a successful, productive QA program. As a research manager, I am very proud of the contributions the QA Team members make to the Division, the Laboratory, and the Agency. I consider it a privilege to be associated with them.

Quality Assurance Project Plans: A Useless Paper Exercise or Valuable Aid ?

Brian A. Schumacher, NERL, ORD, U.S. EPA, Environmental Sciences Division (ESD), Characterization and Monitoring Branch Las Vegas, NV

Two perspectives on the fundamental question "Are quality assurance project plans (QAPPs) a useless paper exercise or a valuable aid?" will be explored. These perspectives include those of a branch chief (i.e., the supervisor/manager) and an active researcher. As a branch chief, when I approach my staff and mention the letters QA, their general autonomic response is a sigh, eyes rolling up to the heavens looking for divine escape possibilities, or the "I knew I should have stayed home today" look. To further exasperate one's staff, tell them that for their new project, they will have to prepare a QAPP or that they need to reexamine and update an existing QAPP. Little do the scientists realize the true value of the QAPP from the managerial perspective. QAPPs are a vital source of information that the branch chief can use for a multitude of different purposes. Research conducted within Characterization and Monitoring Branch (CMB) generally falls into QA category 4 with a few projects falling into QA category 3. Research areas within CMB are diverse and include: improving soil sampling methods with an emphasis on soils contaminated with volatile organic compounds, geophysics, chemometrics, geostatistics, ground water research, technology verification studies, laboratory accreditation, and technology support projects. With this marked diversity of projects, a branch chief can not be an expert in all the different areas. However, through the proper reading and reviewing of the QAPPs, the basic premise(s) of the research project can be learned. This basic knowledge can, in turn, be used during planning exercises as well as during progress and peer reviews. Additionally, the branch chief's review provides the scientist with a different perspective on the project and can help ensure that the research is on the right track. As a research scientist, one of the two greatest benefits to preparing a QAPP (besides getting the boss off your back) is that it gets all your ideas down in writing so that you can remember them, think about them, edit them, and come up with a scientifically sound approach to the research that needs to be accomplished. The other benefit of preparing the QAPP is getting it peer reviewed. The benefit added by getting the opinions and viewpoints of different scientists, not directly involved in the research, is invaluable. It is during the peer review process that ideas the researcher did not originally think about are presented, different approaches are introduced, and questions are asked that make the scientist think about exactly what they are proposing to do. This peer review process, which probably would not occur unless a QAPP was prepared, can only make for a better research project. Thus, to answer the question of "Are QAPPs a useless paper exercise or a valuable aid?", this branch chief/research scientist would have to answer "a valuable aid."

Successful Quality Assurance (QA) Programs in EPA's Office of Research and Development

Brenda T. Culpepper, Director of Quality Assurance & Records Management,
EPA, Research & Development

To implement and maintain a successful QA Program, one must have a partnership and trust among managers, scientists, and QA professionals. Below are the Agency's minimum requirements for a QA program. Additionally, the roles of management, scientists, and the QA professionals are provided.

The EPA Order (5360.1 A2, May 2000) Minimum Requirements for Quality Systems

- QA Manager
- Quality Management Plan
- Sufficient Resources to Implement the Quality System
- Assessments of the Quality System
- Organizational QA Annual Report and Work Plan
- Use of Systematic Work Planning Approach
- Approved QA Project Plans Prior to Any Data Gathering Work or Use
- Assessment of Existing Data, When Used to Support Agency Decisions or Other Secondary Purposes, to Verify They Are of Sufficient and Adequate Quality for Their Intended Use
- QA Requirements for Extramural Agreements
- Corrective Actions Based on Assessment Results
- QA Training for All Levels of Management and Staff

Management

- Ultimate responsible party for implementation of EPA Order
- Authority to make it happen
- Accountable for provision of resources
- Accountable for the Quality Management System

QA Professionals

- QA oversight appropriate for intended use of the data (graded approach)
- More than one right way to be compliant with EPA Order
- "Gotcha" vs. value added
- Assessment of the scientist's ability to reanalyze, reconstruct, defend research
- Thorough record keeping — Just good science
- GLP in a non-GLP laboratory

Scientists

- Responsible for the creation of the QA Project Plan or equivalent
- Verification that QAPP is current and compliant during the lifetime of the project
- Verification of project personnel competence to produce consistent data of the quality required to meet its intended use
- Thorough recordkeeping of all aspects of the study such that the study can be reanalyzed, repeated, or defended
- Independent thinkers (don't like to be told what to do or how to do it)
- First goal — Publish

A successful QA program is one that is part of the culture (i.e., just part of how the research is planned and conducted), not something tacked on after the fact.

Establishing Sensitivity Requirements for Environmental Analyses from Project Data Quality Objectives

Thomas Georgian and Chung-Rei Mao, U.S. Army Corps of Engineers

This article proposes a simple strategy for establishing sensitivity requirements (quantitation limits) for environmental chemical analyses when the primary data quality objective is to determine if a contaminant of concern is greater or less than an action level (e.g., an environmental "cleanup goal," regulatory limit, or risk-based decision limit). The approach assumes that the contaminant concentrations are normally distributed with constant variance (i.e., the variance is not significantly dependent upon concentration near the action level). When the total or "field" portion of the measurement uncertainty can be estimated, the relative uncertainty at the laboratory's quantitation limit can be used to determine requirements for analytical sensitivity. If only the laboratory component of the total uncertainty is known, the approach can be used to identify analytical methods or laboratories that will not satisfy objectives for sensitivity (e.g., when selecting methodology during project planning).

Introduction

There is much confusion in the environmental testing industry concerning how to establish measurement quality objectives (MQOs) for sensitivity from project data quality objectives (DQOs). In part, this stems from debate regarding the best approach for defining and measuring sensitivity. For example, there are many definitions for the term "quantitation limit" but there is no standard approach for determining the lowest concentration at which reported concentrations are considered to be "quantitatively reliable." For example, EPA guidance document QA/G-5 defines the "limit of quantitation" as the "minimum concentration of an analyte or category of analytes in a specific matrix that can be identified and quantified above the method detection limit and within specified limits of precision and bias during routine analytical operations."[1] However, the document does not specify numerical values for precision and bias at the quantitation limit or discuss how to establish project-specific values. As a consequence, the quantitation limit is often defined in a generic manner and sensitivity requirements (e.g., for analytical service contracts) are often established rather arbitrarily, giving rise to data that are not necessarily scientifically defensible. A simple strategy, based upon statistical hypothesis testing, is proposed for establishing MQOs for sensitivity. The approach is applicable when the primary objective is to determine whether the concentration of a contaminant of concern is less than some action level (e.g., an environmental "cleanup" goal, regulatory limit, or risk-based decision limit) and the concentrations are normally distributed with constant variance.

Hypothesis Testing

Statistical hypothesis testing is frequently used to determine whether or not environmental contamination is greater or less than an action level, AL[2]. In hypothesis testing, two mutually exclusive hypotheses are established for some parameter of interest, X. The "alternative hypothesis," H_1, is accepted (i.e., assumed to be true) when a set of measurements provides "strong" or "overwhelming" evidence that indicates that the "null hypothesis," H_0 is false.

Otherwise, the null hypothesis is "accepted" (more accurately, the null hypothesis is not rejected). For example, assume that X denotes the concentration of vinyl chloride in the groundwater of some study area and x denotes the concentration of vinyl chloride reported from the analysis of a particular groundwater sample (i.e., x is some measured value of X). Because of various sources of uncertainty associated with the overall measurement process, repetitive measurements of X will produce a distribution of values. If the distribution is normal and the error is solely random in nature, then the population mean μ represents the "true" value of X and the standard deviation σ represents the "precision" of the overall measurement process. When comparing the measured vinyl chloride concentrations with an action level (e.g., the Maximum Contaminant Level of the Safe Drinking Water Act), one of the following sets of hypotheses could be used:

First Set of Hypotheses:	Second Set of Hypotheses:
H_0: $\mu \geq AL$,	H_0: $\mu \leq AL$,
H_1: $\mu < AL$	H_1: $\mu > AL$

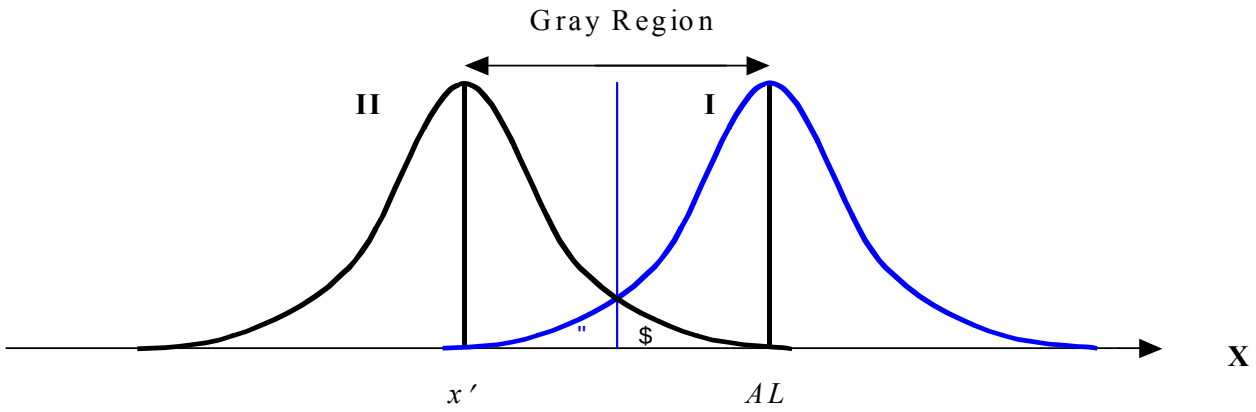
The null hypothesis or "baseline assumption" would be selected (e.g., "$\mu \geq AL$" versus "$\mu \leq AL$") based on the severity of the consequences resulting from its false rejection. For example, a risk assessor may select the null hypothesis "$\mu \geq 2$ µg/L" because erroneously concluding that the groundwater is "clean" (i.e., $\mu < 2$ µg/L) when it is actually "dirty" (i.e., $\mu \geq 2$ µg/L) would be less protective of human health and the environment. The null hypothesis "$\mu \geq 2$ µg/L" would be accepted (e.g., remedial activities would be required) unless the results of the laboratory analyses were to indicate that vinyl chloride is present at a concentration less than 2 µg/L. However, it should be noted that, since H_0 is "$\mu \geq 2$ µg/L," the probability of erroneously concluding that $\mu \geq 2$ µg/L when the groundwater is actually "clean" would be high for concentrations near but less than the action level (e.g., which would result in unnecessary cleanup).

For the first set of hypotheses, it can be shown that

$$AL - x' = \sigma (z_{1-\alpha} + z_{1-\beta}) \qquad (1)$$

The symbol "α" denotes the allowable Type I or "false positive" error (i.e., the probability H_0 will be falsely rejected). The symbol "β" denotes the allowable Type II or "false negative" error, the probability that H_0 will be accepted when H_0 is false (i.e., when $\mu = x' < AL$). The quantity z_p is the $100p^{th}$ percentile of the standard normal distribution (e.g., if $\alpha = 0.05$, then $z_{1-\alpha} = 1.65$). The difference "$AL - x'$" is "width of the gray region" (or "minimum detectable difference"); the quantities x' and AL are the "lower bound" and "upper bound" of the gray region, respectively. The "gray region" consists of the values of X for which the project-specified tolerances for Type I and Type II errors will not be met. The concentration of analyte must be less than x' to demonstrate the groundwater is "clean" at an acceptable level of uncertainty. Note that, if the null hypothesis were "$\mu \leq AL$" (from the second set of hypotheses listed above), the width of the gray region would equal to that in Equation 1, but AL would be the lower rather than upper bound of the gray region.

Figure 1: Upper and lower bounds of the gray region



Establishing Objectives for Sensitivity

The analytical method must be capable of providing quantitative results at concentrations less than or equal to the lower bound of the gray region. Therefore, if H_0 is "$: \geq AL$" and QL denotes the "quantitation limit," the following inequality is a necessary condition for satisfying requirements for sensitivity:

$$QL \leq x' \qquad (H_0: : \geq AL)$$

Note that if H_0 were "$: \leq AL$," then QL would need to be less than or equal to AL.

The quantitation limit, QL, is typically set at a concentration equal to or higher than the lowest initial calibration standard (since instrumental response is usually unknown below the lowest initial calibration standard), and at least three to ten times greater than the detection limit (e.g., the Method Detection Limit defined in 40 CFR, Part 136, Appendix B). The quantitation limit is often defined as 10 σ_L, where σ_L denotes the component of the uncertainty from the "laboratory" or analytical portion of the total uncertainty[3, 4]. Ideally, σ_L accounts for the uncertainty associated with the laboratory analysis (e.g., sample preparation and instrumental analysis) of a homogenized environmental sample. As will be shown, the absolute value of the quantitation limit (e.g., QL = 10 σ_L versus 30 σ_L) is not critical as long as the magnitude of the relative uncertainty at the quantitation limit (σ_L / QL) is acceptably small. However, the uncertainty associated with the laboratory analysis constitutes only a portion of the total uncertainty.

It is reasonable to assume that the laboratory contribution to the total uncertainty, σ^2, is independent of the "field" component of the uncertainty, σ_F^2, so that

$$\sigma^2 = \sigma_L^2 + \sigma_F^2 \qquad (2)$$

The variance for the laboratory portion of the total uncertainty is denoted by σ_L^2 and the variance of the remaining portion of the total uncertainty is denoted as the "field" uncertainty, σ_F^2. Typically, σ_F^2 accounts for uncertainties associated with the sample collection activities in the field and the heterogeneity of the environmental medium being sampled. The variance σ_F^2 can be determined from estimates of σ^2 and σ_L^2. An estimate of σ^2 may be obtained by

implementing some appropriate sampling design for the study area (e.g., a random sampling or stratified sampling scheme as discussed in the USEPA QA-G5S guidance document)[5]. The laboratory portion of the uncertainty, F_L^2, may be estimated from laboratory control sample (or, if available, matrix spike) recoveries[6].

Since the null hypothesis "$: \geq AL$" constitutes the more conservative approach and requires lower quantitation limits, satisfying sensitivity requirements when H_0 is "$: \geq AL$" will be addressed. Since $0 < x' < AL$, it follows that

$$0 < x'/AL < 1 \qquad (3)$$

Dividing Equation 1 by AL and solving for x'/AL gives the result:

$$x'/AL = 1 - (F/AL)(z_{1-\alpha} + z_{1-\beta}) \qquad (4)$$

Substitution of Equation 4 into Equation 3 produces the inequality:

$$0 < F/AL < (z_{1-\alpha} + z_{1-\beta})^{-1} \qquad (5)$$

The quantity F/AL, which will be referred to as the "relative standard uncertainty at the action level," must be less than $(z_{1-\alpha} + z_{1-\beta})^{-1}$ in order to satisfy the allowable decision errors. For example, if $\alpha = \beta = 0.05$, F/AL must be a positive number less than 0.303. (Note that QL approaches zero as F/AL approaches the limiting value of 0.303). For a particular value of F/AL, Equation 4 can be used to calculate the ratio x'/AL, to ensure that $QL/AL \leq x'/AL$ (i.e., $QL \leq x'$). For example, if $F/AL = 0.15$, then QL must be less than or equal to one half the action level to satisfy objectives for sensitivity.

The ratio QL/AL can be expressed as a function of F/QL (rather than F/AL). It follows from Equation 4 that

$$QL/AL \leq 1 - (F/QL)(QL/AL)(z_{1-\alpha} + z_{1-\beta})$$

Solving for QL/AL gives:

$$QL/AL \leq 1 / \{1 + (z_{1-\alpha} + z_{1-\beta})(F/QL)\} \qquad (6)$$

Thus, the required ratio QL/AL can be calculated for various values of the relative uncertainty at the quantitation limit, F/QL.

If the square root of the ratio of the "field" to the "laboratory" components of the uncertainty is denoted by n (i.e., $n = F_F/F_L$), then it follows from Equations 2 and 6 that

$$QL/AL \leq 1 / \{1 + (n^2 + 1)^{1/2} (z_{1-\alpha} + z_{1-\beta})(F_L/QL)\} \qquad (7)$$

Thus, for a given value of the relative standard uncertainty at the quantitation limit, Equation 7 can be used to determine if *QL* is sufficiently small relative to *AL*. To illustrate, *QL*/*AL* is calculated for various values of F_L/QL and *n* in Table 1.

Table 1: *QL*/*AL* for Various Values of F_L/QL and *n* When " = $ = 0.05.

F_L/QL	*QL*/*AL* *n*=0	*QL*/*AL* *n*=3	*QL*/*AL* *n* =5
0.1	0.75	0.49	0.37
0.2	0.60	0.32	0.23
0.3	0.50	0.24	0.17

For example, if $F_L/QL = 0.1$ (i.e., QL = 10 F_L) and $n = 5$, then data quality objectives for sensitivity will not be satisfied unless $QL \leq 0.37\ AL$ (e.g., if $AL = 2$ ppb, then $QL \leq 0.7$ ppb). Note that, as shown in Table 1, *QL*/*AL* decreases as *n* or F_L/QL increases. For a fixed value F_L/QL, the ratio is a maximum when $n = 0$ (i.e., when all of the uncertainty arises from the laboratory portion of the measurement process).

Although it is not realistic to assume that only laboratory measurement uncertainty exists, setting $n = 0$ in Equation 7 is useful for identifying analytical methods or laboratories that will not meet data quality objectives for sensitivity (e.g., when screening potential analytical methods or contract laboratories). For example, if $z_{1-"} + z_{1-\$} = 3.3$, $F_L/QL = 0.30$, and $n = 0$, then $QL/AL \leq 0.50$. Since *QL*/*AL* is greatest when $n = 0$, objectives for sensitivity will not be met for any analytical method for which $QL > 0.5\ AL$, regardless of the magnitude of the uncertainty associated with the field measurement process. Replicate laboratory analyses could be used to reduce the laboratory component of the uncertainty but this would increase the cost. Lastly, it should be noted that it is being assumed that F_L is constant (i.e., does not appreciably vary with concentration) between the *QL* and *AL*, an assumption that will not necessary be valid if $AL >> QL$. In particular, if $F = F_L$ is an increasing function of concentration, Equation 7 will over estimate *QL* (i.e., gives an upper bound limit for *QL*).

When the field as well as the laboratory component of the uncertainty is taken into account, smaller quantitation limits will be required. If F is not precisely known but F_L/QL can be estimated, then some conservative estimate for *n* can be used to calculate *QL*/*AL* using Equation 7. The quantity F_L/QL can be estimated from a laboratory's in-house statistical control limits (e.g., the control limits for the recoveries of laboratory control samples) if the spiking concentrations are near the quantitation limit and the action level. (Note that if F is known and F_L is small relative to F_F, then F will be approximately constant even if F_L is proportional to concentration, and Equation 4 can be used to calculate *QL*/*AL*.) For example, if a laboratory's in-house statistical control range for the recovery of vinyl chloride is 100% ± 60%, then $F_L/QL \approx 0.2$. If $n \approx 5$ (e.g., based upon an investigation of a similar study area), then $QL/AL \leq 0.23$.

Conclusion:

A simple approach to help ensure MQOs for sensitivity will be met when environmental test results are compared with action levels is proposed. The approach utilizes statistical hypothesis testing and assumes that the contaminant concentrations are normally distributed with approximately constant variance (F^2) over the method's calibration range. If the total or "field" portion of the measurement uncertainty can be determined, then the relative uncertainty at the laboratory's quantitation limit can be used to determine if the quantitation limit is sufficiently small relative to the action level. The laboratory component of the total uncertainty provides an upper bound for the required quantitation limit and can be used to identify analytical methods or laboratories that will not satisfy MQOs for sensitivity (e.g., when analytical methodologies are being selected during project planning).

References:

1. USEPA Office of Research and Development, *EPA Guidance for Quality Assurance Plans: EPA QA/G-5;* **1998**.

2. USEPA Office of Research and Development, *Guidance for the Data Quality Objectives Process: EPA QA/G-4;* **1994.**

3. Keith, Lawrence H. *Environmental Sampling and Analysis: A Practical Guide;* Lewis Publishers, Inc.: Michigan, **1991,** p 109.

4. Taylor, John K. *Quality Assurance of Chemical Measurements;* Lewis Publishers, Inc.: Michigan, **1989**, p 79.

5. Office of Environmental Information, *Guidance for Choosing a Sampling Design for Environmental Data Collection: EPA QA/G-5S* (Peer Review Draft); **2000**.

6. Georgian, Thomas, *Environmental Testing and Analysis*. **2000**, 9(6), 20-24.

Getting to the Bottom Line:
Decision Quality vs. Data Quality

Deana M. Crumbling, U.S. EPA Technology Innovation Office

Investigating and restoring contaminated sites faces conflicting goals: site decisions are supposed to be protective and based on sound science, yet costs are to be controlled. Gathering environmental data to support these kinds of decisions can be very expensive. For many years, fixed laboratory analysis for chemical contaminants was the only feasible option, but the expense sharply limited the number of data points that can be generated, compromising the thoroughness of data gathering activities. A culture has developed that focuses a great deal of attention on the quality of chemical analysis, while sample representativeness, generally the largest single source of uncertainty in environmental data, gets less attention.

Recent technology advancements in rapid soil and groundwater sampling tools, field-portable analytical instrumentation, and decision-support software present both opportunity and challenge. Cost-effective generation of real-time data (by whatever means is most feasible) permits the use of the work-flow strategy known commonly as "dynamic work plans," based on real-time decision-making in the field by experienced staff following a pre-approved decision tree. Thoroughly planned and properly implemented, dynamic work plans have been shown to save about 30 to 50% of project costs due to fewer remobilizations to fill data gaps and more efficient use of expensive equipment and labor (such as backhoes and drill rigs). Dynamic work plans have also resulted in more thorough site characterization because immediate feedback allows data gaps and problems to be resolved in real-time. It is now possible to manage the critical sampling uncertainties that stem from matrix heterogeneity. It is always more cost-effective in the long-term to make the right decision the first time.

Despite obvious benefits, acceptance is growing slowly; many institutional barriers remain. For example, field methods are often dismissed as "field screening" and are not used to their full potential. This is largely because many practitioners find it difficult to access the appropriate technical expertise needed to design sampling and analytical plans that generate data of known and documented quality, while being simultaneously driven by project decision quality. This presentation discusses how to distinguish analytical quality from data quality, and then to link data quality firmly to data use and decision quality. This approach creates a framework for using field analytical methods as highly cost-effective tools that produce much higher decision quality than possible under the current paradigm.

INTRODUCTION

Exhortations for "sound science" and "better quality data" within the context of regulatory decision-making are increasingly popular. Is the current data quality model sufficient to achieve sound science? Is "data quality" really the key issue, or is there something more fundamental at stake? Although this paper

focuses primarily on contaminated site cleanup, many of these issues are broadly applicable to other areas of environmental management.

Since 1979, U.S. Environmental Protection Agency (EPA) policy has required an Agency-wide quality system, with the goal of providing "environmental data of adequate quality and usability for their intended purpose of supporting Agency decisions." (Ref. 1). Yet the linkage between data quality and data usability for decision-making is easily lost from programmatic and project planning and implementation. "Data quality" is too often viewed as some independent standard established by outside arbiters. Project managers tend to follow a checklist of accepted methods as a primary means of achieving data quality. Striving for "high quality data" per this pervasive data quality model has proven to be an expensive and sometimes counterproductive exercise.

"Sound science" in regulatory and project decision-making is achieved by managing decision uncertainty. Acceptable "data quality" is achieved by managing data uncertainty, so that the data can support the aspect of decision-making for which it is intended. Managing uncertainty, either of decisions or of data, requires careful planning and appropriate technical skills. Calls for "sound science" and "better data quality" are meaningless without simultaneous commitment to improve programmatic and project planning using scientifically qualified staff. If policy-makers desire to see sound science used in environmental decisions, they need to provide a coherent vision that can steer development of infrastructure that will support management of decision quality at the project level. Data quality IS important, but without a guiding vision that focuses on decision quality, free-floating mandates for "data quality" drain already scarce resources.

It is a mistake to assume that scientific data are (or can be) the only basis for regulatory decision-making. Science may be able to provide information about the nature and likelihood of consequences stemming from an action, but the decision to pursue or reject that action (i.e., accept or reject the risk of consequences) based on scientific information is within the province of values, not science. Even the choice of how much uncertainty is tolerable in statistical hypothesis testing is in the realm of values. Thus, it is appropriate that many non-scientific considerations feed into a regulatory decision-making process. This does not invalidate a foundation of "sound science" as long as the various roles of science and values are differentiated, and any underlying assumptions and other uncertainties in both data and decision-making are openly declared with an understanding of how decision-making would be affected if the assumptions are erroneous.

DECISION QUALITY AS DEFENSIBILITY

The term "decision quality" implies that decisions are defensible (in the broadest sense). Ideally, decision quality would be equivalent to the correctness of a decision, but in the environmental field, decision correctness is often unknown (and perhaps unknowable) at the time of decision-making. When knowledge is limited, decision quality hinges on whether the decision can be defended against reasonable challenge in whatever venue it is contested, be it scientific, legal, or otherwise. Scientific defensibility requires that conclusions drawn from scientific data do not extrapolate beyond the available evidence. If scientific evidence is insufficient or conflicting, decision defensibility may rest properly on other considerations. No matter what those considerations are, "defensibility" implies there is honest and open acknowledgment of the full range of uncertainties impacting the decision-making process.

Managing scientific defensibility is extremely difficult when the science behind a new initiative is immature. This was undeniably the situation when Superfund and other site cleanup programs were created in the 1980's. In a classic chicken-and-egg dilemma, fledgling waste programs were asked to create procedures despite the fact that the scientific and technical foundation for such programs barely existed. At the same time, programs were called upon to legally defend their cleanup decisions. To develop the needed scientific theory, practice, and tools for measuring and mitigating contamination and its effects, the government began to pour funding

into research to understand the complex relationships among environmental, chemical, and health phenomena. Despite the immaturity of the science, policy-makers and the public expected that cleanup activities would begin and proceed immediately. Few anticipated the daunting technical complexities that would be encountered by cleanup programs as they leapt into this unknown sphere of science and engineering.

FIRST-GENERATION STEPPING STONES THAT BECAME STUMBLING BLOCKS

When immediate action is desired, but knowledge and expertise are not yet sufficient to plot the smartest plan of attack, the most reasonable tactic is to initially create a consistent, process-driven strategy based on the best available information that everyone can follow while experience and knowledge accumulates. Certainly this made sense for the emerging cleanup programs. To be consistent with sound science, however, such a process-driven approach should be openly acknowledged by all participants as the first approximation that it is, with the understanding that one-size-fits-all oversimplifications will be discarded in favor of specific performance goals as more scientifically sound information becomes available. Although science is comfortable viewing first approximations as short-lived stepping stones subject to revision, this view is much less welcome in an litigious regulatory atmosphere. Furthermore, as individual cleanup programs proliferated at the state and local levels, first approximations become more and more solidified in a bureaucratic process that naturally prefers predictability and consistency. The net result is that the regulatory and procedural infrastructures that support project implementation have trouble keeping up behind the maturing science.

This lag manifests in various ways. A prime example is the prevailing concept of "data quality" as applied to environmental analytical chemistry data. Advances in characterization technologies and strategies (such as using field analytical methods to manage sampling uncertainties and support real-time decision-making in the field) risk rejection simply because they do not fit the ancestral data quality model. This is true despite their potential to save time and money, while simultaneously improving decision quality (Ref. 2-4).

The data quality model used in site cleanup programs was a first approximation based on incomplete knowledge of environmental systems. At its root are several assumptions about analytical chemistry test results:

1) "Data quality" is determined by the accuracy and documentation of the chemical analysis procedure (traditionally performed in a laboratory).
2) The accuracy of analyses on environmental samples can be ensured by consistently performing all analyses according to strictly prescriptive regulator-approved methods.
3) Analytical uncertainty (i.e., the degree to which the accuracy of the analytical results are in question) can be managed according to a checklist regimen of quality control procedures that rely largely on ideal matrices such as reagent water or clean sand to establish method performance.
4) Laboratory quality assurance is equivalent to, and substitutable for, project quality assurance.
5) With "cook book" analytical procedures for the laboratory, and a list of approved analytical methods in hand for project planning, the need for environmental analytical chemistry expertise can be minimized in the environmental laboratory and eliminated from project planning.

Decision-makers accepted these tacit assumptions at the birth of cleanup programs, although scientists warned of questionable validity (Ref. 5, 6). This oversimplified "analytical quality is data quality" model supported the imperative to "define the nature and extent of contamination" as a first approximation toward a sampling and analysis strategy for hazardous waste sites. It was hoped that "defining the nature and extent" would produce information (data) that would tell the project manager what to so with the site. Naturally, it was impossible in the early days to predict the kind of cleanup and land reuse decisions that

would be faced later on, so each site had to be a "study." Data had to be collected without any chance of predicting in advance exactly how the data might be used to support decisions later on in the project. This unfocused approach can work if there are sufficient resources to repeatedly return to the site to fill newly discovered data gaps as piece-meal identification of site decisions and their associated uncertainties proceeds. Although the best strategy for that time, advancing knowledge and 20 years of program experience means that this process is out-of-date. In fact, as program budgets shrink and rapid reuse of sites is desired, it is no longer viable. "Defining the nature and extent" without first identifying project goals amounts to groping around in the dark. It carries a serious danger that decision uncertainties will not be identified in a timely manner, so data collection designs will be inadequate to support scientifically sound decision-making. If there are not sufficient funds to continue data collection until decision uncertainties are managed, there is a strong incentive to downplay or ignore decision uncertainties. This in turn increases the chance that decision errors could pose unacceptable risks to receptors, and/or will waste resources through ineffective remedial actions and costly repetition of effort to fill data gaps that contribute to decision errors (Ref. 7, 8). This is the antithesis of sound science.

EVOLVING A SECOND-GENERATION DATA QUALITY MODEL

To set the stage for an updated data quality model, we must clarify the term "data quality." According to EPA's Office of Environmental Information, data quality is "the totality of features and characteristics of data that bear on its ability to meet the stated or implied needs and expectations of the user/customer" (Ref. 9). What data users need, ultimately, is to make the correct decisions. Therefore, data quality cannot be viewed as some arbitrary standard, but must be judged according to its ability to supply information that is representative of the decision that the data user intends to make. Said in a different way, anything that compromises data representativeness compromises data quality, and data quality cannot be assessed except in relation to the intended decision (Ref. 10).

The assumptions of the current data generation model fail this test. The root cause of this failure is the fact that the data used to make project decisions are generated from environmental samples (i.e., specimens) that are drawn from highly variable and complex parent matrices (such as soils, waste piles, sludges, sediments, groundwater, surface water, waste waters, soil gas, fugitive airborne emissions, etc.). This fact has several repercussions:

1) The concept of representativeness demands that the scale (spatial, temporal, chemical species, etc.) of supporting data be the same (within tolerable uncertainty bounds) as the scale of the intended decisions. In contaminated site projects, the true state (such as the concentrations of contaminants across space or time or the properties of the matrix that control contaminant fate and transport) can easily vary markedly over small or large scales. High variability at one scale may be inconsequential if viewed over a different scale. It is not resource-feasible to characterize the "true state" of all relevant properties of the site at all possible scales. So, without project planning that is anchored in first understanding the scale over which decision-making will occur, selecting the scale over which to "define nature and extent" is guesswork.

2) The concept of representativeness can be broken into sample representativeness and analytical representativeness, both of which are critical to managing data uncertainties:

    - Sample representativeness includes specimen selection, collection, preservation, and subsampling procedures. All are crucial to data quality, but the representativeness of specimens is difficult to ensure without sufficient sampling density to understand the scale of matrix heterogeneity. Because of the highly heterogeneous nature of many environmental matrices, most of the uncertainty in data stems from the sampling side of data generation (Ref. 3).

    - Analytical representativeness involves selecting an analytical method that produces data that is representative of the decision. Causes of analytical non-representativeness include selecting the wrong method or erroneously interpreting method results, and failing to recognize when matrix

interferences degrade method performance to the point where erroneous decisions would be made. If such interference is found, sound science demands that method modification or an alternate method be used to compensate. Evaluating analytical performance on ideal matrices (reagent water and clean sand) provides little reassurance that equivalent performance is being achieved on project-specific samples.

3) The wide range of decisions, contaminants, matrices and interferences encountered in site cleanup programs and the pace of technology development make it impossible for prescriptive analytical requirements to accommodate the multitude of complex and interacting variables that determine method performance. Regulatory flexibility for the selection and operation of analytical methods is not only vital to ensuring representative results, but also fosters acceptance of highly cost-effective, second-generation technologies and strategies.

4) The scientific and technical complexities of site cleanup require that appropriate scientific expertise be involved in up-front project planning to identify decision goals and to design data collection strategies, in design implementation, and in data interpretation. Without appropriate expertise, identification and management of relevant uncertainties does not occur, data quality is frequently mismatched to data use, and sound science is not achieved.

5) Arbitrary regulatory requirements for "data quality" should be avoided since this short-circuits sound science. Regulations should focus on requirements for performance that demonstrate explicit management of decision uncertainty.

6) Conceivably there will be circumstances where it is more cost-effective to manage decision uncertainty by simply choosing the most protective action without generating data.

A second-generation data quality model for the environmental field will explicitly recognize that

- Data quality is an emergent property arising from interaction between the attributes of the data (such as its bias, precision, sensitivity, and other characteristics that together contribute to data uncertainty) and the intended use of the data (assist with the management of decision uncertainty).
- Data uncertainty is comprised of both sampling and analytical uncertainties.
- Analytical uncertainty in a test result arises from the analytical uncertainty of the measurement method itself and from interaction between the sample matrix and the analytical process. The analytical uncertainty arising from the method itself is only a fraction (and often a negligibly small fraction) of the overall data uncertainty. The impact of sample matrix on analytical uncertainty varies to a greater or lesser degree depending on how well the analytical methodologies have been matched to the sample matrix and to the data needs.
- Sampling uncertainty accounts for the majority (sometimes nearly all) of the data uncertainty. This uncertainty is managed by increasing the sampling density and/or by targeting sample collection designs to yield the most valuable information. Sample representativeness requires that all aspects of sampling design be matched to the scale of decision-making.
- Procedures to estimate and report data uncertainties (e.g., uncertainty intervals) to the data user need to be developed and followed.
- Appropriate technical expertise is required to implement this model.

SUMMARY

Years of experience with investigating and cleaning contaminated sites have made it clear that data quality cannot be managed independent of the overarching goal of decision uncertainty management. Pursuing arbitrary notions of "data quality" becomes an elusive, aimless, disconnected resource sink that fails to achieve sound science. Data quality (management of data uncertainty) and decision quality

(management of decision uncertainty) are distinctly different endeavors, both of which are critical to the pursuit of "sound science." Yet their roles are easily confounded in the regulatory arena. Isolated attempts to address data quality issues that fail to recognize and address fundamental conflicts between outdated models and contemporary scientific knowledge only perpetuate problems. Pursuing policies based on sound science will challenge government agencies to modernize first-generation environmental models and static regulatory strategies to accommodate the ever-evolving progressive nature of science itself.

REFERENCES

1. U.S. Environmental Protection Agency. 2000a. EPA Order 5360.1 A2, Policy and Program Requirements for the Mandatory Agency-wide Quality System. Washington, DC.
2. Crumbling, D.M. 2001. Current Perspectives in Site Remediation and Monitoring: Using the Triad Approach to Improve the Cost-Effectiveness of Hazardous Waste Site Cleanups. EPA 542-R-01-016. August. Issue paper available at http://cluin.org/tiopersp/
3. Crumbling, D.M. 2001. Applying the Concept of Effective Data to Environmental Analyses for Contaminated Sites. EPA 542-R-01-013. August. Issue paper available at http://cluin.org/tiopersp/
4. Crumbling, D.M., C. Groenjes, B. Lesnik, K. Lynch, J. Shockley, J. van Ee, R.A. Howe, L.H. Keith, and J. McKenna. 2001. Managing Uncertainty in Environmental Decisions: Applying the Concept of Effective Data at Contaminated Sites Could Reduce Costs and Improve Cleanups. Environmental Science & Technology 35:9, pp. 404A-409A. Article reprint is available at the following website: http://cluin.org/download/char/oct01est.pdf
5. Fairless, B.J. and D.I. Bates. 1989. Estimating the quality of environmental data. Pollution Engineering March:108-111.
6. Homsher, M.T.; F. Haeberer; P.J. Marsden; R.K. Mitchum; D. Neptune; and J. Warren. 1991. Performance Based Criteria, A Panel Discussion. Environmental Lab, October/November. Articles available at http://cluin.org/download/char/dataquality/perfbased.pdf
7. Francoeur, Thomas L. 1997. Quality Control: The Great Myth. Proceedings of the Field Analytical Methods for Hazardous Wastes and Toxic Chemicals Conference, January 29-31, 1997, Las Vegas, NV, Air & Waste Management Association, Pittsburgh, PA. Article available at: http://cluin.org/download/char/dataquality/qc_greatmyth.pdf
8. Popek, E.P. 1997. "Investigation versus Remediation: Perception and Reality" in Proceedings of WTQA'97—the 13th Annual Waste Testing and Quality Assurance Symposium, pp. 183-188. Paper available at http://cluin.org/products/dataquality/
9. U.S. Environmental Protection Agency (USEPA). 2000. OEI Quality System 2000: Office of Environmental Information Management System for Quality. http://www.epa.gov/oei/quality.htm
10. U.S. Environmental Protection Agency (USEPA). 2000. Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9 QA00 Update). EPA 600/R-96/084. July. http://www.epa.gov/quality/qs-docs/g9-final.pdf

Performance-Based Approach and Data Quality

Ben Hull, U.S. EPA Office of Radiation and Indoor Air
John Griggs, U.S. EPA, Office of Radiation and Indoor Air, NAREL

The Multi-Agency Radiological Laboratory Analytical Protocols (MARLAP) Manual is a document which provides guidance for the planning, implementation and assessment phases of those projects which require the laboratory analysis of radionuclides. MARLAP's basic goal is to provide guidance and a framework for project planners, managers and laboratory personnel to ensure that radioanalytical laboratory data will meet a project's or program's data requirements and needs. To attain this goal, the manual seeks to provide the necessary guidance for national consistency in radioanalytical work in the form of a performance-based approach for meeting a project's data requirements. The guidance in MARLAP is designed to help ensure the generation of radioanalytical data of known quality appropriate for its intended use.

MARLAP was developed by a workgroup which included representatives from the Environmental Protection Agency (EPA), the Department of Energy (DOE), the Department of Defense (DOD), the Nuclear Regulatory Commission (NRC), the National Institute of Standards and Technology (NIST), the U.S. Geological Survey (USGS), the U.S. Food and Drug Administration (FDA). State participation in the development of the manual involved contributions from representatives from the Commonwealth of Kentucky and the State of California. The draft manual is currently available at the follow website: http://www.eml.doe.gov/marlap/. The manual is undergoing an EPA Science Advisory Board review in 2002.

MARLAP

The MARLAP Manual is a document which provides guidance for the planning, implementation and assessment phases of those projects which require the laboratory analysis of radionuclides. MARLAP's basic goal is to provide guidance and a framework for project planners, managers and laboratory personnel to ensure that radioanalytical laboratory data will meet a project's or program's data requirements and needs. To attain this goal, the manual seeks to provide the necessary guidance for national consistency in radioanalytical work in the form of a performance-based approach for meeting a project's data requirements. The guidance in MARLAP is designed to help ensure the generation of radioanalytical data of known quality appropriate for its intended use.

MARLAP was developed by a working group which included representatives from the Environmental Protection Agency; the Department of Energy; the Department of Defense; the Nuclear Regulatory Commission; the National Institute of Standards and Technology; the U.S. Geological Survey; the U.S. Food and Drug Administration; the Commonwealth of Kentucky; and the State of California. Since MARLAP employs a performance-based approach to laboratory measurements, the guidance contained in the manual is applicable to a wide range of programs, projects and activities which require radioanalytical laboratory measurements. Examples of data collection activities that MARLAP supports include site characterization; site cleanup and compliance demonstration; decommissioning of nuclear facilities; remedial and removal actions; effluent monitoring of licensed facilities; environmental site monitoring; background studies and waste management activities.

MARLAP is divided into two main parts. Part I is aimed primarily at project planners and managers and provides guidance on project planning with emphasis on analytical planning issues and analytical data

requirements. Part I also provides guidance on developing project-specific analytical requirements, selecting analytical methods, preparing project plan documents and radioanalytical statements of work (SOWs), obtaining and evaluating radioanalytical laboratory services, data verification, data validation, and data quality assessment.

Part II of MARLAP is aimed primarily at laboratory personnel and provides guidance in the relevant areas of radioanalytical laboratory work. The chapters in Part II are intended to serve as a resource base of information on the laboratory analysis of radionuclides and provide guidance on a variety of activities performed at radioanalytical laboratories including sample preparation; sample dissolution; chemical separations; instrument measurements; data reduction, etc. Part II also has chapters on measurement statistics, laboratory quality assurance and quality control and waste management for radioanalytical laboratories. While the chapters in Part II do not contain detailed step-by-step instructions on how to perform certain laboratory tasks, the chapters do provide information on many of the options available for these tasks and discuss advantages and disadvantages of each.

Performance-Based Approach

MARLAP provides the necessary guidance for using a performance-based approach to meet a project's analytical data requirements. In a performance-based approach, the project-specific analytical data requirements that are determined during directed planning serve as measurement performance criteria for analytical selections and decisions. The project-specific analytical data requirements also are used for the initial, ongoing, and final evaluation of the laboratory's performance and the laboratory's data. MARLAP provides guidance for using a performance-based approach for all three phases, planning, implementation and assessment, of the data life cycle for those projects that require radioanalytical laboratory data. This involves not only using a performance-based approach for selecting an analytical protocol, but also using a performance-based approach for other project activities, such as developing acceptance criteria for laboratory quality control samples, laboratory evaluations, data verification, data validation, and data quality assessment.

There are three major steps or processes associated with a performance-based approach. The first is clearly and accurately defining the analytical data requirements for the project. The second involves using an organized, interactive process for selecting or developing analytical protocols to meet the specified analytical data requirements and for demonstrating the protocol's ability to meet the analytical data requirements. The last major activity involves using the analytical data requirements as measurement performance criteria for the ongoing and final evaluation of the laboratory data, which would include data verification, data validation, and data quality assessment. MARLAP provides guidance in all three of these areas. Within the constraints of other factors, such as cost, a performance-based approach allows for the use of any analytical protocol that meets the project's analytical data requirements. For all relevant project activities, the common theme of a performance-based approach is the use of project-specific analytical data requirements that are developed during project planning and serve as measurement performance criteria for selections, evaluations, and decision-making.

Performance Objectives: Data Quality Objectives and Measurement Quality Objectives

One of the outputs of a directed planning process is DQOs for a project or program. DQOs are qualitative and quantitative statements that clarify the study objectives; define the most appropriate type of data to collect; determine the most appropriate conditions from which to collect the data; and specify tolerable limits on decision error rates (ASTM D5792; EPA, 2000). DQOs apply to all data collection activities associated with a project or program, including sampling and analysis. In particular, DQOs should encompass the "total uncertainty" resulting from all data collection activities, including analytical and sampling activities. From an analytical perspective, a process of developing the analytical data requirements from the DQOs of a project is essential. These analytical data requirements serve as measurement performance criteria or objectives of the analytical process. MARLAP refers to these performance objectives as "measurement quality objectives" (MQOs). The MARLAP Manual provides

guidance on developing the MQOs from the overall project DQOs (Chapter 3). MQOs can be viewed as the analytical portion of the DQOs and are therefore project-specific. MARLAP provides guidance on developing MQOs during project planning for select method performance characteristics, such as method uncertainty at a specified concentration; detection capability; quantification capability; specificity, or the capability of the method to measure the analyte of concern in the presence of interferences; range; ruggedness, etc. An MQO is a statement of a performance objective or requirement for a particular method performance characteristic. Like DQOs, MQOs can be quantitative and qualitative statements. An example of a quantitative MQO would be a statement of a required method uncertainty at a specified radionuclide concentration, such as the action level—i.e., "a method uncertainty of 3.7 Bq/kg (0.10 pCi/g) or less is required at the action level of 37 Bq/kg (1.0 pCi/g)." An example of a qualitative MQO would be a statement of the required specificity of the analytical protocol—the ability to analyze for the radionuclide of concern given the presence of interferences—i.e., "the protocol must be able to quantify the amount of ^{226}Ra present given high levels of ^{235}U in the samples." The MQOs serve as measurement performance criteria for the selection or development of analytical protocols and for the initial evaluation of the analytical protocols. Once the analytical protocols have been selected and evaluated, the MQOs serve as criteria for the ongoing and final evaluation of the laboratory data, including data verification, data validation, and data quality assessment. In a performance-based approach, analytical protocols are either selected or rejected for a particular project, to a large measure, based on their ability or inability to achieve the stated MQOs. Once selected, the performance of the analytical protocols is evaluated using the project-specific MQOs. The MARLAP manual is currently available at the follow website: http://www.eml.doe.gov/marlap/. The manual is undergoing an EPA Science Advisory Board review in 2002.

Quantifying Uncertainty: Are We There Yet?

Louis Blume, U.S. Environmental Protection Agency, Great Lakes National Program Office, 77 West Jackson Boulevard, Chicago, Illinois 60604
Judy Schofield and Ken Miller, DynCorp I&ET, Inc., 6101 Stevenson Avenue, Alexandria, VA 22304

EPA's Great Lakes National Program Office (GLNPO) is leading one of the most extensive studies of a lake ecosystem ever undertaken. The Lake Michigan Mass Balance Study (LMMB Study) is a coordinated effort among state, federal, and academic scientists to monitor tributary and atmospheric pollutant loads, develop source inventories of toxic substances, and evaluate the fate and effects of these pollutants in Lake Michigan. A key objective of the LMMB Study is to construct a mass balance model for several important contaminants in the environment: PCBs, atrazine, mercury and trans-nonachlor. The mathematical mass balance models will provide a state-of-the-art tool for evaluating management scenarios and options for control of toxics in Lake Michigan.

At the outset of the LMMB Study, managers recognized that the data gathered and the model developed from the study would be used extensively by data users responsible for making environmental, economic, and policy decisions. Environmental measurements are never true values and always contain some level of uncertainty. Decision makers, therefore, must recognize and be sufficiently comfortable with the uncertainty associated with data on which their decisions are based. The quality of data gathered in the LMMB was defined, controlled, and assessed through a variety of quality assurance (QA) activities, including QA program planning, development of QA project plans, implementation of a QA workgroup, training, data verification, and implementation of a standardized data reporting format. As part of this QA program, GLNPO has been developing quantitative assessments that define data quality at the data set level. GLNPO also is developing approaches to derive estimated concentration ranges (interval estimates) for specific field sample results (single study results) based on uncertainty. The interval estimates must be used with consideration to their derivation and the types of variability that are and are not included in the interval.

The Great Lakes, which contain 20% of the world's freshwater, are a globally important natural resource that are currently threatened by multiple stressors. While significant progress has been made to improve the quality of the lakes, pollutant loads from point, non-point, atmospheric, and legacy sources continue to impair ecosystem functions and limit the attainability of designated uses of these resources. The U.S. EPA's Great Lakes National Program Office (GLNPO) instituted the Lake Michigan Mass Balance Study (LMMB Study) to measure and model the concentrations of representative pollutants within important compartments of the Lake Michigan ecosystem. The LMMB Study was a coordinated effort among Federal, State, and academic scientists to monitor tributary and atmospheric pollutant loads, develop source inventories of toxic substances, and evaluate the fates and effects of these pollutants in Lake Michigan. A key objective of the LMMB Study is to construct a mass balance model for several important contaminants in the environment: PCBs, atrazine, mercury and trans-nonachlor. The mathematical mass balance models will provide a state-of-the-art tool for evaluating management scenarios and options for control of toxics in Lake Michigan.

At the outset of the LMMB Study, managers recognized that the data gathered and the model developed from the study would be used extensively by data users responsible for making environmental, economic,

and policy decisions. Environmental measurements are never true values and always contain some level of uncertainty. To address these issues, GLNPO employed a comprehensive suite of tools to define, control, and assess data quality. These tools included QA program planning, development of QA project plans by each of the Principle Investigators (PIs) responsible for collecting and/or analyzing samples in the LMMB, implementation of a QA workgroup to guide and monitor QA activities, up-front and ongoing training, independent verification of all field and laboratory results, and development and implementation of a standardized data reporting format. As part of this QA program, GLNPO also has been applying statistical approaches to develop quantitative assessments of data quality.

For the LMMB Study, all data were categorized, reported, and evaluated by "focus groups". A focus group is defined by sample medium (e.g., lake water, tributary water, fish, plankton, sediment, etc), by pollutant type (e.g., PCBs, atrazine, mercury, and trans-nonachlor), and by the PI responsible for analyzing the samples. Data quality assessments were conducted by focus group to reflect the distinct sampling and analytical procedures used by each PI. GLNPO is conducting quantitative assessments for each data focus group for six data quality attributes: sensitivity, system precision, analytical precision, system bias, analytical bias, and a novel attribute, percent variability due to sampling and analytical measurement uncertainty. These assessments were presented in a paper title, Will Anyone Ever Read the Lake Michigan Mass Balance Quality Assurance Report, at EPA's 20th Annual Conference on Managing Environmental Quality Systems. These quantitative estimates reflect data quality at the focus group level.

Deriving an Interval Estimate

Study modelers have requested an interval estimate for single study results, expressed as a range of concentrations based on uncertainty, to set initial conditions and test model outputs. Ideally, such intervals would be derived from collection and analysis of repeated replicates of a given sampling unit. Due to resource constraints, repeated replicates of a given sampling unit were not generally collected and analyzed for the LMMB study, so GLNPO has been exploring alternate approaches that involve use of quality control (QC) sample results to derive these interval estimates.

The types of QC samples available to derive interval estimates varies according to focus group, because the LMMB Study was a performance-based study in which PIs were afforded a great deal of flexibility in choosing the QC tools that would be employed to meet study objectives. The estimates that can be derived, and the variability associated with that estimate, depend on the QC data available for a given focus. Examples of QC samples and the variability included in an interval estimate based on the these QC samples are provided in Table 1. Consideration of the variability that is and is not included in each type of interval estimate is critical when interpreting these estimates.

Table 1. QC samples and the variability included in an interval estimate derived from these results

QC sample	Description	Variability included in interval estimate

Field Spiked Sample (FSF)	A routine field sample that is split in the field. This split is fortified in the field with known concentration of analyte and analyzed in the lab. The other split is analyzed without fortification.	All field and analytical activities including: sampling; sample shipment, storage and handling; and analysis including sample matrix effects.
Field Control Sample (FCM)	Aliquot of reagent water or other neutral item (resin, sand) to which known quantity of target analyte is added in the field. Otherwise handled, transported, and analyzed same as RFS.	All field and analytical activities including: sampling; sample shipment, storage and handling; and analysis without sample matrix effects.
Laboratory Spiked Sample (LSF)	A routine field sample that is split in the lab. This split is fortified in the lab with known concentration of analyte and analyzed in the lab. The other split is analyzed without fortification.	All analytical activities and sample matrix effects.
Laboratory Control Sample (LCM)	Aliquot of reagent water or other neutral item (resin, sand) to which known quantity of target analyte is added. Contains same reagents, solvents, standards, etc. as RFS.	All analytical activities without sample matrix effects.

Recovery-based Interval Estimate

One approach to developing interval estimates is used in several of the 1600-series methods developed by the EPA Office of Water as a means by which laboratories should monitor their performance. The approach, which is described in OW's Guidance on the Documentation and Evaluation of Trace Metals Data Collected for Clean Water Act Compliance Monitoring (EPA-821-B-96-004, July 1996), uses the mean recoveries and standard deviation of ongoing QC measurements (i.e., spiked field samples, spiked reagent water, standard reference materials, or surrogate spikes) to establish confidence bounds around analytical results. The interval is estimated as:

$$\text{Mean recovery} \pm (\text{Standard deviation} * t)$$

where:
Mean and standard deviation are the mean and standard deviation of all QC sample recoveries, and t is the 97.5th percentile of the student's t distribution with n-1 degrees of freedom, where n is the number of QC sample results.

The recovery-based interval estimate can be used to estimate the true value of a reported result and to construct bounds around the result. For example, if the result reported is 10 ppb and the recovery-based interval estimate is 84% +/- 25% (i.e., the mean recovery is 84% and the standard deviation of the recovery times the t statistic is 25%) then the true value will be in the range of 9.2 - 16.9 ppb with 95% confidence. This range is derived as follows:

$$\text{Lower limit} = [10/(0.84 + 0.25)] = [10/1.09] = 9.2 \text{ ppb}$$
$$\text{Upper limit} = [10/(0.84 - 0.25)] = [10/0.59] = 16.9 \text{ ppb}$$

The recovery-based interval estimate provides variability information for sampling and the analytical activities associated with the sample result, depending on the type of QC sample used to calculate the interval. If the interval is based on spiked reference matrix samples (as opposed to spiked field samples), matrix effects and associated variability will not be reflected in the estimate. Ideally, the interval estimates would be the same based on spiked field samples and spiked reference matrix samples. Any difference

could be attributable to random error or sample matrix effects. Deriving an interval using LCM samples will likely result in a tighter interval than one calculated using spiked field samples because sources of variability regarding matrix effects are not reflected in the estimate. This is shown in Table 2. For illustration interval estimates are applied to the median of field sample results for a given focus. Because median results are adjusted based on mean recovery of QC sample results, the resulting interval estimates are not necessarily centered around the median value.

Table 2. Comparison of Recovery Intervals based on Different Types of Spiked Data

Focus	Median Result	Spike Type	# QC results	Interval	
				Percent Recovery	Concentration
Fish Mercury	101 ng/g	Laboratory Spiked Field Sample	9	69.64% to 115.9%	87.13 to 145.0 ng/g
		Laboratory Reference Sample	24	87.74% to 112.4%	89.89 to 115.1 ng/g
Atmospheric Total Phosphorus	3.6 µg/L	Laboratory Spiked Field Sample	53	82% to 126%	2.9 to 4.4 µg/L
		Laboratory High Check (reference matrix standard)	162	84% to 113%	3.2 to 4.3 µg/L

Duplicate-based Interval Estimate

An interval also can be derived based on the variability between the field sample (RFS) results and their associated field duplicate (FD) results (within-pair variance). The within pair variance is estimated as:

$$s^2_w = MSE = \frac{1}{n} \sum_{i=1,\ldots,n} s^2_{(RFS_i,\ FD_i)}$$

where: n is the number of routine field sample/field duplicate pairs,
$s^2_{(RFS\,i,\ FD\,i)}$ is the variance between the routine field sample and field duplicate in pair i.

The interval is derived using the standard deviation (the square root of S^2_w) as follows:

$$\text{Exp}\{\log(\text{Sample result}) \pm (s_{L,w} * t)\}$$

where: $s_{L,w}$ is the square root of $S^2_{L,w}$, calculated based on log-transformed RFS and field duplicate results, and t is the 97.5th percentile of the student's t distribution with n degrees of freedom, where n is the number of duplicate pairs.

The field sample and field duplicate results are log-transformed prior to calculating the pair variances to

reflect the fact that absolute variability of analytical data increases with increasing concentration. Log transformation also may address the skewed distribution often observed in field data. Because the recovery-based interval estimate is percentage-based, and because the distribution of recoveries is more likely to follow a normal distribution than a distribution of field results, log transformation is not necessary for a recovery-based interval. For some LMMB focuses, a small number of field duplicates were collected and analyzed compared to a large number of field sample results collected at a large number of sites. Therefore, the interval estimate may not accurately reflect the variability for all field samples. The duplicate-based interval estimate is more valid for focuses where there are a large number of field duplicates taken at a large number of study stations.

The interval estimate also can be adjusted for bias by dividing the sample result by the mean recovery of other QC sample results (such as those used to create the recovery-based interval estimate) prior to log-transformation. This adjustment reflects components of bias based on the QC sample used as presented in Table 1 and does not include *variability* associated with the bias estimate.

Table 3 provides examples of recovery-based interval estimates and duplicate-based interval estimates for several example focus groups. As in Table 2, for illustration the interval estimate is applied to the median of field sample results for a given focus. In addition, these median results were adjusted based on the mean recovery of QC sample results for that focus, as described in the paragraph above, for both interval types, therefore, the resulting intervals are not necessarily centered around the median value.

Table 3. Examples of Interval Estimates for Single Study Results

Focus	Median Result	Interval Type	QC Type	# QC results	Interval	Width
Tributary Mercury	4.5 ng/L	Recovery-based	Lab Spiked Sample	53	3.5 to 5.8 ng/L	2.3 ng/L
		Duplicate-based	Field Duplicate	46	3.0 to 6.5 ng/L	3.5 ng/L
Open Lake Mercury	0.3 ng/L	Recovery-based	Laboratory Performance Check	68	0.20 to 0.40 ng/L	0.20 ng/L
		Duplicate-based	Field Duplicate	13	0.17 to 0.42 ng/L	0.25 ng/L
			Laboratory Duplicate	68	0.19 to 0.38 ng/L	0.19 ng/L
Tributary Ortho-phosphate	0.0086 mg/L	Recovery-based	Lab Spiked Sample	59	0.00828 to 0.00980 mg/L	0.00152 mg/L
		Duplicate-based	Field Duplicate	19	0.00335 to 0.0241 mg/L	0.0207 mg/L

The interval estimates based on field QC sample results are usually wider than those based on laboratory QC sample results. This occurs, in part, because the intervals based on field QC samples include variability associated with sample collection and handling as well as analytical activities, whereas the interval estimates based on laboratory QC samples reflect variability associated with the analytical activities only. These interval estimates are confidence intervals for a single known study result. For some applications, a prediction interval may be more appropriate.

The interval estimates must be used with consideration to their derivation and the types of variability that are and are not included in the interval. The available QC data for each focus will likely dictate potential approaches for estimates. The *Recovery-based Interval Estimate* is one of the simplest approaches and likely will be one of the most broadly applied for LMMB data (i.e., most focuses have laboratory QC data that can be used to construct this interval). However, the variability reflected in the interval depends on the type of QC sample used to derive the interval. Current efforts are focusing on potential approaches to combine imprecision and bias variability into a single interval estimate, in order to encompass all available uncertainty information based on the QC results for a given focus group.

The Role of Quality Assurance in the Evaluation of Two Landfill Bioreactor Operational Techniques at an Existing Landfill

J.T. Markwiese, A.M. Vega, R. Green, P. Black

A Quality Assurance Project Plan (QAPP) was prepared to document the primary objectives and the data collection and interpretation efforts for two landfill bioreactor studies at the Outer Loop Landfill in Louisville, Kentucky, operated by Waste Management, Inc. The multiyear bioreactor studies include facultative landfill bioreactor and staged aerobic-anaerobic landfill bioreactor demonstrations. Treatment and control groups were established and consist of separate and distinct landfill units; each unit is composed of paired cells. The primary objective for both studies is to evaluate the treatment effect on waste stabilization and settlement relative to the controls.

Besides describing the planned experimental design and data analysis aspects of the project, the QAPP also includes details regarding ensuring sample representiveness and analytical quality assurance procedures. All measurements needed to evaluate the primary objectives (critical measurements) are supplemented by quality control checks, including auditing procedures, so that data of known quality are generated. These quality activities will help ensure that the data generated are appropriate for their intended use. Namely, to provide the landfill community with potential alternatives for rapid and controlled reduction of the waste mass in a landfill containment system. The approach described here has potential application to other landfill facilities across the country.

The Outer Loop Landfill operated by Waste Management has been used for waste disposal for approximately 35 years. Two multi-year studies are proposed for the Outer Loop Landfill, including a Facultative Landfill Bioreactor (FLB) Study, and an Aerobic-Anaerobic Landfill Bioreactor (AALB) Study. Bioreactor landfills are designed to accelerate the biological stabilization of landfilled waste by means of leachate recirculation that enhances decomposition by creating a favorable environment for microbially-mediated waste stabilization. Enhanced waste stabilization should reduce the potential for future environmental problems because waste stabilization occurs within the operating life of the liner. In addition, bioreactor technology can reduce long-term requirements for monitoring gas migration and cover maintenance while minimizing the time required for profitable energy production through gas recovery. The effectiveness of bioreactor technology is being jointly evaluated by EPA and Waste Management, Inc., through a 5-year Cooperative Research and Development Agreement. In the Outer Loop study, treatment and control groups consist of separate and distinctlandfill units; each unit is composed of two paired cells. The FLB study is being performed in paired landfill cells that are generally 4-6 years of age covering approximately 47 acres and the AALB study is being performed in paired one-year old landfill cells covering 12 acres. A separate unit of paired cells containing approximately 2-3 year old waste is used as the control for the FLB and AALB studies. Because landfill units are filled sequentially (placement of waste in a particular cell is only initiated after the current waste-receiving cell is completely filled), individual units in this study are not directly comparable with respect to time. However, the control cells will provide an adequate treatment reference by considering them as temporally offset from the treatment cells. For example, consider the comparison between FLB cells and the control. As mentioned, FLB waste is generally 4-6 years old and control waste is about 2-3 years old. In three years, control waste will be approximately the same age as present-day FLB

waste. Therefore, control samples collected three years following the initiation of the FLB treatment will represent the FLB cells as they were when leachate was first introduced.

Facultative Landfill Bioreactor (FLB) Study The primary objective is to evaluate waste stabilization and settlement resulting from nitrate-enriched leachate application to test cells relative to waste stabilization in control cells. This approach is based on two premises: (1) the addition of leachate will moisten and promote degradation of the waste and (2) microorganisms present in the landfill waste will use nitrate in the leachate as a terminal electron acceptor for anaerobic metabolism. As nitrate containing liquid moves through the upper sections of the FLB, denitrifying bacteria convert nitrate to dinitrogen gas. This transformation of nitrate-nitrogen to gaseous nitrogen should result in a net loss of nitrogen from the landfill and enhanced waste degradation under anaerobic conditions. Enhanced waste degradation is expected because, relative to other terminal electron acceptors in anaerobic environments (e.g., sulfate and carbon dioxide), nitrate offers the greatest thermodynamic yield per unit carbon respired; i.e., there is relatively more energy available for bacterial growth. Enhanced microbial growth is expected to be equivalent to enhanced waste stabilization.

Aerobic-Anaerobic Landfill Bioreactor (AALB) Study The primary objective is to evaluate waste-stabilization enhancement resulting from the sequential establishment of aerobic and anaerobic conditions in the AALB cells relative to waste stabilization in the control cells. Waste is treated aerobically, similar to composting, by injecting air into the waste for approximately 45 days. After aeration is discontinued, the waste is moistened with liquids, and anaerobic conditions are quickly established. The rationale behind this sequential approach is to promote the rapid decomposition of food waste and other easily degradable organic matter in the aerobic stage of treatment with the intent of reducing the amount of fermentable organic matter entering the anaerobic stage. This could shorten the acid generating phase of anaerobic waste decomposition and result in a more rapid onset of methanogenesis

Critical Measurements Landfilled waste typically progresses through five phases of degradation, including: (1) adjustment or acclimation; (2) transition; (3) acidogenesis; (4) methanogenesis; and (5) maturation. This degradation process can be collectively considered as waste stabilization. At any given time, landfill cells may be characterized as experiencing one of the above phases. But because waste is deposited in a landfill cell over time (months to years), waste-stabilization phases tend to overlap and sharp boundaries between phases are not typical. It is expected, however, that the bioreactor treatments will increase the rate of transition through the various phases relative to the control. It is further expected that this enhanced transition to stabilized waste will be discernable with trend analyses. The critical measurements (*italicized*) employed in this study were selected to capture aspects of waste stabilization over time.

Acclimation. During acclimation, microbial populations are in a state of adjustment and respiration rates are generally low. *Waste moisture* tends to increase and available *oxygen* is slowly consumed during this phase. Since the atmospheric-oxygen supply to the buried waste is diffusion limited, the concentration of oxygen in the landfill cell begins to decrease.

Transition. In the transition phase, conditions turn anaerobic as the oxygen consumption rate increases due to metabolism of readily degradable wastes. Complex organic matter is broken into simpler forms (e.g., organic acids) and energy that is not captured by cells during respiration is given off as heat. *Waste* and *leachate temperature* concomitantly increase during organic-matter degradation. Other respiration by-products (*carbon dioxide* and *volatile organic acids*) begin to increase in leachate.

Acidogenesis. During acidogenesis the accumulation of volatile organic acids reaches its peak due to metabolism and fermentation of organic matter. The increase in *chemical oxygen demand* and *biochemical oxygen demand* indirectly reflects this increase in degradable metabolites. In addition, the

high concentration of acids increases hydrogen ion activity, reflected by decreased *waste* and *leachate pH*. In the near absence of oxygen, metabolism shifts to anaerobic bacteria capable of utilizing alternate electron acceptors (e.g., nitrate and sulfate).

Methanogenesis. In the methanogenic phase, the supply of most electron acceptors is exhausted. Methanogenic bacteria ferment organic acids to *methane* and *carbon dioxide* while other methanogens utilize CO_2 as their terminal electron acceptor. Consequently, *gas* (methane and CO_2) *volume* and production rates increase. Anaerobic respiration is a proton-consuming process and this is reflected by an increase in pH values in the waste and leachate.

Maturation. The maturation phase represents the end-point of landfill *stabilization (surface GPS measurements)*. The overall conversion of complex wastes to leachable organic acids (phases 2 and 3) and gaseous products (phase 4) also serves to reduce the waste volume and *organic solids* and to increase *waste density*. Maturation occurs when degradable organic matter, and consequently microbial growth, is limited. This is reflected by decreases in the biochemical methane potential and gaseous metabolic by-products methane and CO_2. Concentrations of organics in leachate remains steady but at substantially reduced levels relative to earlier phases.

The terms "settlement" and "stabilization" are used interchangeably in this project. As noted above, settlement or stabilization is largely defined in terms of microbial processes because stabilization is dependent upon microbial breakdown of organic matter. In addition to the biological and chemical parameters listed, settlement of the test and control cells will be measured by a professional surveying team by taking quarterly readings of 40 to 80 global positioning system points in each treatment. The critical measurements listed above (and the critical measurement biochemical methane potential) directly support the primary project objective of evaluating waste stabilization.

Data Evaluation Given the difference in age between the treatment and control landfill cells and the limited number of cells available for the investigation, robust statistical methods will be employed. Typically non-parametric methods are more robust than parametric ones, hence they are recommended here. Comparability of treatment and control data (i.e., comparability among landfill cells) will be carefully examined before performing any statistical analyses. If the treatment and control data resulting from this project are determined to be incomparable, the recommendations and conclusions will focus on the weight of evidence provided by exploratory data analysis to evaluate the effectiveness of the treatment. These techniques include calculation of summary statistics and investigation of the data using pictures and graphs.

Assuming the data from treatment and control are comparable, the Mann-Kendall test for trend will be employed in time series analyses. This test uses the relationship between time-adjacent results to determine whether there is sufficient evidence to detect an increasing or decreasing trend. The assumption is that the treatment and control will follow the same trend for a given measure, but the time period over which the trend occurs may be different, with the trend in the treatment cell being accelerated over time compared with the control cell. In this case, the differences between treatment and control will get larger over time, hence the differences will show an increasing trend, even if seasonal fluctuations are present.

Quality Assurance and Quality Control To optimally generate known-quality data, a scientifically sound and strictly followed quality control program must be incorporated into the sample collection and analytical aspects of the project. Relative to other solid matrices (e.g., soil) landfill waste is extremely heterogeneous. Several parameters were considered in developing a sampling strategy to represent the chemical, biological and physical status of a landfill in the most representative way possible. Because each cell's leachate drains to a central sump, samples collected at sumps should be representative of the entire cell. Systematic locations for the gas extraction wells and waste boring locations were chosen to maximize the coverage within the zone of maximum vertical resolution (i.e., away from the sides of the cell). The gas collected from multiple collection points is also mixed and this helps ensure

representativeness of gas data. To minimize solid-waste sample variability, large sample volumes are being collected and analyzed in each treatment. GPS measurements are also being used to assess stabilization and thus no samples are necessary. Matrices will be sampled to provide a "snapshot" of the historical contents of the landfill. The goal is to effectively choose enough points on the landfill to get a complete picture upon combining the information from each snapshot.

The QA objectives defined in the analytical program for the critical measurements are summarized in Table 1 in terms of the following data quality indicators: precision, accuracy, method detection limits, and completeness. Comparability and representativeness are achieved by the use of standard EPA methods or well-documented SOPs throughout the duration of the project and through adherence to a well-defined sampling strategy for capturing adequate samples to characterize properties of each matrix. If necessary, reanalysis of the samples will be conducted when possible. Corrective actions (detailed in the QAPP) taken in response to non-compliant data will be documented and summarized in the project's final report and the impact on project objectives will be evaluated and discussed.

Table 1. Quality Assurance Objectives for Critical Measurements

Measurement	Matrix	Time Point*	Precision a	Accuracy b	RDLs c	Units	Com-plete-ness
Chemical Oxygen Demand	Leachate	G	± 20%	100 ± 20%	5	mg/L	100%
Biochemical Oxygen Demand	Leachate	G	± 20%	100 ± 30%	2	mg/L	100%
Leachate temperature (d)	Leachate	FE	± 1°C	± 1°C	N/A	°F	100%
pH	Leachate	FE	± 0.1	± 0.1	N/A	-log H+	100%
Volatile Org Acids	Leachate	G	± 20%	100 ± 25%	0.1	mg/L	100%
Waste Temperature (d)	MSW	FE	± 1°C	± 1°C	N/A	°F	100%
Waste Settlement (e)	MSW	TP	± 5 cm	± 5 cm	N/A	cm	100%
Organic Solids (f)	MSW	G	± 25%	± 0.1%	N/A	%	100%
Moisture Content (f)	MSW	G	± 2%	± 0.1%	N/A	%	100%
pH (g)	MSW	G	± 0.1	± 0.1	N/A	-log H+	100%
Biochemical Methane Potential	MSW	G	± 20%	100 ± 20%	1	ml/g	100%
Waste Density	MSW	G	N/A	(j)	N/A	lb/cubic yard	100%
CH_4, CO_2, O_2 (h)	Gas	G	(h)	(h)	Appendix C	% (vol)	100%
Gas Volume (i)	Gas	G	± 5%	100 ± 5%	N/A	Ft^3	100%

* Samples are collected as a grab(G), field electrode (FE) or time point (TP) at the point of collection. GPS measures represent unique temporal/spatial sampling points.

Precision expressed as the relative percent difference (RPD) between spiked duplicates and/or lab duplicates (biochemical methane potential precision assessed with the relative standard deviation [RSD] of triplicate samples)

(a) Accuracy expressed as the % recovery of matrix spikes or as the measurement of a known standard

(b) RDLs are the reporting detection limits as devised by the lowest calibration standard or weight.

(c) Precision and accuracy objectives for temperature are based upon thermocouple specifications

(d) Precision and accuracy objectives for GPS are based upon manufacturer specifications (Trimble model 4800), positioning accuracy determination outlined in QAPP Section 4.4.3.

(e) Precision and accuracy objectives for moisture and organic solids are based upon calibration requirements for analytical balances and duplicate weight measures of the same sample.

(f) Accuracy for pH is based upon known standards. Precision is based on sample duplicate readings.

(g) Gas composition precision (sample duplicate) and accuracy (certified gas standard) are as follows: methane and carbon dioxide precision, ± 10% (RPD), accuracy, 100 ± 10%; oxygen precision 30% (RPD), accuracy 30%.

(h) Gas volume precision and accuracy are based upon manufacturer specifications and factory certification of the flow meter used.

(i) Scale is calibrated monthly and must be accurate to ± 1% of true weight.

Data Reporting, Data Reduction and Data Validation For analytical data to be scientifically valid, defensible, and comparable, the correct equations and procedures must be used to prepare the data. Evaluation of measurements is a systematic review process to provide assurance that the data are adequate for their intended use. The process includes the following activities:

- Auditing measurement system calibration and calibration verification;
- Auditing QC activities;
- Screening data sets for outliers;
- Reviewing data for technical credibility vs. the sample site setting;
- Checking intermediate calculations; and
- Certifying the above process.

Lab data validation procedures are required to employ an independent analyst to review all aspects of data generation, including the calculation steps used to generate sample concentrations. Outer Loop subcontracted laboratories will conduct this activity as part of their normal operations. Individual analysts will review the data generated each day to determine the need for corrective action or rework. Data will also undergo a second review process conducted by one of three independent reviewers (typically a second analyst, a lab manager or a QA manager). In addition, EPA or Waste Management will perform data validation separate from that performed by the laboratories on at least 10% of all data.

Audits Audits are an independent means of confirming the operation or capability of a measurement system, and of independently documenting the use of QC measures designed to generate valid data of known and acceptable quality. For all tests/methods conducted by laboratories, the results for the performance evaluation (PE) samples received and processed by the laboratories (just prior to, during, and immediately following their involvement in the project) for purposes of compliance with laboratory certification requirements relating to these analyses (or where the laboratory is not regulated, PE samples submitted blind to analysts by laboratory management) will be provided to Waste Management. For all failed PE results the laboratory will institute remedial actions and where valid performance of the measurement system cannot be established, the laboratory will establish corrective actions. These corrective actions will include evaluation of testing data that may have been affected, notification of the client if project data may have been affected, and amended reports with data appropriately qualified if and when the laboratory determines that data have been affected. A system audit is a qualitative determination of the overall ability of a measurement system to produce data of known and acceptable quality, by an evaluation of all procedures, personnel, equipment, etc., utilized to generate the data. It is an evaluation of whether adequate QC measures, policies, protocols, safeguards, and instructions are inherent in the measurement system to enable valid data generation and subsequent actions. EPA QA personnel have conducted the first of biannual (every two years) technical systems audits for field trials and the resulting laboratory evaluations of the samples collected during field testing.

Summary In summary, the planning process outlined here will help ensure that data of known quality are interpretable and useful for assessing treatment effects associated with bioreactor technology. This approach helps ensure the defensibility of decisions regarding the efficacy of bioreactor technology. The experiences and results of performing a systematic planning operation for the Outer Loop Landfill will potentially be used in conducting, evaluating and regulating bioreactor landfills in the future.

Working on Quality Improvement Case Study: Sampling and Analysis Plans Under the Navy Installation Restoration Program

Narciso A. Ancog, QA Officer, Naval Facilities Engineering Command, Southwest Division
John Dirgo, Ron Ohta, and Greg Swanson, Tetra Tech EM Inc.

This paper summarizes the joint efforts of Naval Facilities Engineering Command, Southwest Division (SWDIV), and Tetra Tech EM Inc. (Tetra Tech) quality assurance (QA) personnel to improve the quality of environmental data collected under the SWDIV Installation Restoration Program. Improvement efforts, which began in November 1999 and are ongoing, have focused on three main areas: document format; management and communications; and use of information technology. Our results are presented as a case study, identifying the challenges faced by the QA organizations at SWDIV and Tetra Tech, the approach we developed to meet each challenge, and the lessons we learned in the process.

BACKGROUND

The Naval Facilities Engineering Command, Southwest Division (SWDIV), based in San Diego, California, manages the environmental restoration activities at Naval installations in the Southwest, including the states of California, Arizona, New Mexico, and Nevada. SWDIV implements an environmental Quality Assurance Program (QAP) to respond to quality requirements established by various regulatory agencies. These agencies include, but are not limited to, the U.S. Environmental Protection Agency (EPA) Region IX, the California Department of Toxic Substances Control, and the California State Water Resources Control Board. The SWDIV QAP supports the Installation Restoration Program by specifying procedures that ensure the usability and defensibility of chemical data. The SWDIV Quality Assurance (QA) officer implements the QAP and is responsible for developing and enforcing Environmental Work Instructions (EWI) (internal quality procedures), reviewing and approving Sampling and Analysis Plans (SAP), and field surveillance of sampling efforts to identify quality issues. Quality managers from various contractors work with the SWDIV QA officer to ensure that all activities involving sampling and analysis meet SWDIV and contractual requirements for quality.

Tetra Tech EM Inc. (Tetra Tech) has been a SWDIV contractor since November 1, 1999, when management of Tetra Tech's Comprehensive Long-term Environmental Action Navy (CLEAN) contract was transferred from the Engineering Field Activity West in the San Francisco Bay area to SWDIV in San Diego. Under its current and past CLEAN contracts, Tetra Tech has performed hazardous waste site characterization and remediation work since 1989 at more than 20 U.S. Navy installations in California. This work is supported by project teams operating out of 10 offices, and typically requires Tetra Tech to prepare 30 to 40 SAPs over each 6-month period.

Shortly before contract management was transferred, SWDIV and Tetra Tech QA personnel met to discuss how operating under the SWDIV QAP would affect planning and execution of environmental data collection efforts. Much of the discussion focused on SWDIV's requirements and expectations for SAPs as key planning documents in environmental data collection. The dialogue that began in this initial meeting has continued over the last 22 years and has resulted in several improvements in the SAPs developed by SWDIV and its contractors.

This paper is a case study of how we implemented these improvements and the lessons we learned. We also discuss factors that helped us make improvements and how barriers that hindered progress were overcome.

CHALLENGES AND SOLUTIONS

This section presents seven different challenges SWDIV and Tetra Tech faced on the road to improved SAPs. We identify the key technical and organizational factors that contributed to each challenge, the approaches we took to meet each challenge, and the results we achieved.

Challenge 1: Making the Transition to SWDIV Management

The transition of Tetra Tech's CLEAN contract to SWDIV presented significant challenges for both organizations. For example, Tetra Tech had developed an installation-wide quality assurance project plan (QAPP) for many installations. When a new task order that required environmental data collection was assigned, project teams typically prepared an abbreviated plan that was supported by the installation-wide QAPP. However, a full SAP was required for each new investigation under SWDIV management. Tetra Tech had been working on the CLEAN contract for 10 years when the transition occurred, and there was a significant amount of inertia to overcome. SWDIV remedial project managers (RPM) were eager to make progress on cleanups at their newly assigned installations, and Tetra Tech project teams were equally eager to please their new clients. In some cases, RPMs set and project teams agreed to aggressive schedules that allowed little time for QA personnel to review SAPs.

Approach: Tetra Tech and SWDIV QA personnel spent considerable effort during the first 6 months after the transition in establishing communication procedures and educating project teams about SWDIV requirements for SAPs. Tetra Tech posted SWDIV's EWIs on its Intranet so that they would be available as guidance to project teams. Tetra Tech had already routinely assigned an installation coordinator (IC) as a focal point for all work at each Navy installation. QA personnel used the established monthly IC conference calls and IC meetings (held every 6 months) to communicate the importance of meeting SWDIV QA requirements.

The strong position of SWDIV's QA organization helped facilitate the transition. The SWDIV QA officer is part of a review board and has a direct impact on Tetra Tech's performance evaluations. As a result, Tetra Tech's program and project managers become more sensitive to quality issues, which increases the effectiveness of Tetra Tech's QA organization. The SWDIV QA officer also provided Tetra Tech with detailed feedback, via a Document Review Evaluation Form, for all SAPs that were submitted. The form provided summary quality ratings as well as specific comments on each SAP.

Results: By the end of the initial 6-month period (November 1999 through May 2000), Tetra Tech project teams were aware of most SWDIV QA requirements. Tetra Tech had initiated a tracking system so that both SWDIV and Tetra Tech QA personnel were informed of upcoming SAPs and could better schedule reviews. Communication procedures between the QA and project organizations for both SWDIV and Tetra Tech had been worked out, but only after some difficult situations had been resolved. Detailed SAP review comments from the SWDIV QA officer identified several possible targets for quality improvement and provided the basis for the next three challenges.

Challenge 2: Integrating Field Sampling and Quality Assurance Project Plans

The SWDIV QA officer provided Tetra Tech with detailed feedback for each of the 36 SAPs that were submitted during the first 6 months after the transition of the CLEAN contract to SWDIV management. Each SAP consisted of two separate documents C a field sampling plan (FSP) and a QAPP. One recurring problem that the SWDIV QA officer identified was inconsistency between the FSP and the QAPP.

Approach: Tetra Tech questioned the concept that a SAP must consist of a separate FSP and

QAPP and recommended integrating the two into one consolidated plan to avoid redundancy and eliminate inconsistencies. Tetra Tech noted that the QAPP framework specified in *EPA Requirements for Quality Assurance Project Plans (EPA QA/R-5)* might be appropriate for the integrated SAP. In particular, QAPP elements B1 through B3, which focus primarily on sampling, could be used to incorporate elements of the FSP into an integrated SAP.

In July 2000, Tetra Tech and SWDIV QA personnel discussed the pros and cons of an integrated SAP. Based on these discussions, Tetra Tech developed a detailed outline for the integrated SAP that includes all the elements specified in EPA's QA/R-5 requirements and is nearly identical in organization. Tetra Tech also identified an upcoming project that could use the integrated SAP format on a trial basis. Concurrently, the SWDIV QA officer sought approval from the regulatory agencies to use the integrated SAP format.

Tetra Tech submitted its first integrated SAP in September 2000. The document received favorable reviews within Tetra Tech, SWDIV, and the regulatory agencies. In October 2000, the SWDIV QA officer approved the integrated SAP format for all subsequent SAPs under Tetra Tech's CLEAN contract.

Results: Continued use of the integrated SAP has demonstrated that this format is well accepted by both project teams and regulators. The integrated document offers several advantages over two separate documents. It is shorter; requires less effort to prepare; provides a single, comprehensive guide to all project personnel; and avoids problems associated with cross-referencing between a separate FSP and QAPP. As a result, SWDIV issued a directive in September 200l that all of its contractors should begin using the integrated SAP to promote consistency among contractors and across installations.

Challenge 3: Addressing Recurring Deficiencies in SAPs

In addition to inconsistencies between FSPs and QAPPs, the SWDIV QA officer identified a number of other deficiencies in the 36 SAPs that were submitted between November 1999 and May 2000. By reviewing the detailed comments for these documents and grouping the comments into categories, Tetra Tech was able to identify several common problems. The categories included sampling procedures; analytical methods; data quality objectives (DQO); roles and responsibilities; references to guidance; sample handling and custody; and document format.

Approach: In September 2000, Tetra Tech proposed developing a standard template, based on the integrated SAP format discussed above, as a tool to ensure consistency and to address common deficiencies identified by the SWDIV QA officer. The template would include standard, recommended language for sections of the document that do not change significantly from one project to the next (for example, sample handling and custody, reports to management, and data verification and validation methods). The template would also build in corrections to other problems, such as incorrect citations of SWDIV EWIs and other guidance, and incorrect descriptions of SWDIV's roles and responsibilities. Finally, each section of the template would include instructions to indicate language that was considered standard, identify project-specific content that was needed, and provide examples of well-written, project-specific content (such as data quality objectives, sampling process design, and sampling methods).

Tetra Tech developed a draft template, starting from an integrated SAP that had been reviewed and approved by both SWDIV and the regulatory agencies. In April 2001, Tetra Tech and SWDIV QA personnel reviewed the organization and content of the draft template. Tetra Tech revised the draft template to incorporate several of SWDIV's suggestions for improvement, and in May 2001, the SWDIV QA officer approved the final template for use.

Results: After the final template was released, the SWDIV QA officer identified fewer deficiencies and noted that the quality of Tetra Tech's SAPs had improved significantly. The template has allowed Tetra Tech's project teams to reduce the time needed to prepare a SAP and

to focus resources on project-specific issues. Finally, because much of the template consists of standard language, review time for Tetra Tech and SWDIV QA personnel is also reduced.

However, the template is not fool proof. Recently (in December 2001), the SWDIV QA officer noticed a decline in the quality of the SAPs and pointed this out to Tetra Tech. Thus, we have learned that we need to work hard to maintain improvements. With new personnel always coming into the CLEAN program, a key part of this maintenance is facilitating good communications between QA personnel and project teams. In addition, continuous improvements to the template are needed to address new issues as they arise.

Challenge 4: Developing Consistent Data Quality Objectives

From time to time, the SWDIV QA officer has noted that DQOs were not developed in a consistent manner in Tetra Tech's SAPs. In particular, the decisions to be made (Step 2) and decision rules (Step 5) were not consistent in some SAPs. Other common problems have included unclear discussions of the tolerable limits on decision errors (Step 6) as well as confusion over judgmental versus statistical sampling designs (Step 7).

Approach: Tetra Tech recognized that the DQO process is essential to developing a sampling design that is consistent with the project objectives and decisions to be made. Tetra Tech also recognized that the DQO process requires a multidisciplinary understanding and complex judgments. As a result, we focused on training to further the understanding of our project teams on the DQO process. Initial training, which took place in 1999 and 2000, is being supplemented by refresher training in 2002. Tetra Tech has also identified a small group of personnel with extensive experience in developing DQOs and DQO-based sampling designs. Personnel from this group are available to provide support to project teams that request assistance.

Results: Training and internal DQO support have helped improve consistency in developing and presenting DQOs in Tetra Tech's SAPs. However, continued training and support are needed to sustain the gains and to keep project teams informed of the recent DQO developments and tools.

Challenge 5: Modifying SAPs

During 2001, several SWDIV RPMs and Tetra Tech project teams expressed the need for a process to modify SAPs without rewriting and resubmitting the entire document. The SWDIV QA officer also noted that some Tetra Tech project teams were resubmitting complete SAPs for projects where only relatively minor changes in the field investigation were contemplated, an inefficient use of resources.

Approach: SWDIV already had in place a procedure for revising and amending SAPs (EWI #2, Review, Approval, Revision, and Amendment of Sampling and Analysis Plans). However, it was clear that neither the Tetra Tech project teams nor the SWDIV RPMs thoroughly understood this procedure. Tetra Tech and SWDIV QA personnel therefore worked with project teams and RPMs to increase their awareness and understanding of EWI #2 and the circumstances under which amendments to the SAP were appropriate. Tetra Tech developed a standard format for amendments, and disseminated that information at IC meetings and through Intranet postings. This allowed project teams across installations to quickly and consistently prepare amendments.

Tetra Tech also worked with the SWDIV QA officer to develop and disseminate a protocol for the appropriate use of field change orders. With this protocol in place, project teams have three choices for documenting changes to a SAP:

Submit a complete new SAP when major changes are planned or when the original SAP is outdated or inadequate. Examples of major changes are revisions to the DQOs or the sampling design.

Prepare an amendment to the SAP when the changes involve numbers of samples, sampling locations, or related topics, but the project objectives and sampling design remain

unchanged.

Prepare a field change order when the modification is minor and is the result of unanticipated field conditions.

SWDIV and Tetra Tech QA personnel review and approve all SAPs, amendments, and field change orders to ensure that project teams are using these options appropriately. Further, we encourage project teams to work with Tetra Tech and SWDIV QA personnel to determine the most appropriate option to use and to resolve any project-specific planning issues prior to writing the document.

Results: Amendments to SAPs are now being processed routinely, allowing Tetra Tech project teams to document changes to the SAP efficiently while maintaining appropriate QA oversight. Tetra Tech has only recently distributed guidance for using field change orders to project teams. However, preliminary indications are that field change orders will be an effective tool for documenting minor modifications to SAPs.

Challenge 6: Integrating Multiple Contractors

With multiple contractors working concurrently on multiple Navy installations within the SWDIV arena, an effective means of communicating and sharing ideas and experiences is essential to ensure consistency among contractors and consistent implementation of the SWDIV QAP.

Approach: SWDIV established a Data Quality Council (DQC) that is made up of the quality managers from all major SWDIV Installation Restoration Program contractors. The DQC meets quarterly with the SWDIV QA officer to discuss quality issues, review SWDIV requirements, discuss technical topics, and share quality improvement ideas.

Results: The DQC is an effective forum for contractors to share experiences and ideas for quality improvement. Council members have provided practical input to SWDIV EWIs, which SWDIV updated in November 2001. Members of the DQC have also shared ideas for developing abbreviated SAPs for field investigations that are limited in scope, electronic review and approval of SAPs, and data management.

Challenge 7: Expediting Review and Approval of SAPs

Many installations that SWDIV manages have established aggressive schedules with the regulatory agencies for completing cleanups. These expedited schedules are particularly true for installations that are part of the Base Realignment and Closure (BRAC) programCSWDIV faces intense political and public pressure to clean up BRAC installations and return them to civilian use. Aggressive schedules mean that there is often very little time for QA review and approval of SAPs. Complicating this is the fact that QA reviewers and project teams are often located in different offices. Even overnight delivery of documents is sometimes not fast enough to meet project schedules at some installations. This challenge is particularly difficult for both the SWDIV and Tetra Tech QA organizations because we need to maintain quality standards and requirements while avoiding the perception that we are impeding the progress of planned field investigations.

Approach: To address this challenge, Tetra Tech has implemented an electronic system for internal review of SAPs. When draft SAPs are ready for QA review, project managers e-mail the documents to the reviewer or post the documents on a file transfer protocol (ftp) site for download. Electronic review comments are returned by e-mail or via the ftp site to the project manager and project team for resolution. Revised documents are returned electronically for rapid re-review and approval. SWDIV is considering the use of electronic review for selected documents, as well as electronic signature by the QA officer, to further expedite review and approval of SAPs.

***Results*:** Electronic review of SAPs has shortened the time between preparation and implementation and has helped SWDIV RPMs and Tetra Tech project teams meet aggressive schedules. Other advantages of electronic review include reduction in paper waste, elimination of errors in transcribing edits suggested by reviewers, and the ability to provide clearly written comments and editorial suggestions.

CONCLUSIONS

Over the past 22 years, Tetra Tech and SWDIV have implemented several changes to improve the quality of SAPs and to more effectively meet SWDIV QA requirements. These quality improvements have focused on three main areas:

Document formats: Replacing separate FSPs and QAPPs with an integrated SAP and using a standard SAP template have reduced technical inconsistencies, resulted in more concise SAPs, and condensed document preparation and review time. Implementing SWDIV's procedures (EWI #2) for amendments and associated technical direction for field change orders has allowed changes to SAPs to be appropriately documented.

Management and communications: Detailed written feedback from the SWDIV QA officer was critical in demonstrating the need to improve SAPs and in targeting specific areas for improvement. SWDIV's Data Quality Council has created a forum for the SWDIV QA officer and major contractors to bring up and resolve technical and administrative issues that directly affect implementation of SWDIV's QAP. Internal communication tools, such as Tetra Tech's Intranet and regular IC meetings, have helped increase awareness of and compliance with SWDIV QA requirements among Tetra Tech project teams.

Information technology: Internal electronic review of SAPs allows work at multiple contractor offices and multiple installations to proceed efficiently.

Two other factors have been critical in driving the quality improvement efforts described in this paper. The first factor is the strong role that the QA officer plays within the SWDIV organization, which helps ensure that Tetra Tech program management pays attention to quality-related issues. The SWDIV QA officer's evaluation of Tetra Tech's performance directly affects Tetra Tech's award fee on the CLEAN contract. The second factor is the cooperative relationship that has been built between SWDIV and Tetra Tech personnel. This relationship has evolved over time, helped us understand each other's perspectives, permitted us to resolve difficult issues in a collegial manner, and provided a forum for continuous improvement.

Adventures in Environmental Data Reporting: High Tech, Low Tech and Everything In Between

or

Wisconsin DNR's Reporting Systems Move Toward the Future

Donalea Dinsmore, Wisconsin DNR

Wisconsin Department of Natural Resources (DNR) began looking at data reporting systems even before we started our Quality Management Plan. Historically, the data we received was little more than the facility identification, sample designation (e.g. Outfall 001, well #25b), the results of what was measured, and an indication of the laboratory performing the analyses. Even with that limited information, with thousands of regulated facilities, data entry is a huge job. To add to the burden, PCS, EPA's database has had limited compatibility with other databases so data may need to be transcribed again before it reaches EPA. We are very interested in streamlining the process in a manner that preserves data integrity.

Electronic reporting systems have relied on diskettes accompanied by a paper submittal with a signature certifying the accuracy of the information. Air or hazardous waste emissions are reported through a Consolidated Reporting System, a computer software program. DNR sends facility-specific information to over 5,000 facilities each year. Reports may be returned via the Internet or diskettes through the mail.

DNR is working on additional options for reporting data. Last year, DNR updated its system for transmitting data from our State Lab of Hygiene. We plan to expand this portal to accept data from commercial laboratories. Our initial efforts will focus on Safe Drinking Water compliance data. In addition, we are completing a pilot using data entry on web-based forms for DMRs. If we secure funding, we anticipate that machine-to-machine transfer may be available as soon as next year.

Three initiates may shape our data reporting future significantly:

- *Comprehensive data standards that include data quality indicators*
- *NR 148 - Data Reporting Rule*
- *e-Government initiative - WI-MAP Authentication*

In the past two years, DNR developed data standards that cover chemicals in multiple media as well as flora and fauna. We are developing a strategy to implement these standards department-wide. These standards specify what we believe are the minimum elements necessary to establish the context for the data collection effort and a means to evaluate the level of quality.

At many regulated facilities, the same staff handles environmental monitoring and reporting for various compliance programs. The different rules, language, and reporting conventions used in the compliance program leads to confusion and complexity in reporting the monitoring data.This fall, we begun working on a data reporting rule, NR 148, with the following goals:

- *Consolidate compliance reporting requirements into a single Code*
- *Standardize reporting conventions where possible*
- *Specify the minimum required data elements, implement data standards*
- *Identify performance criteria for electronic reporting*

We see this as an opportunity to optimize our reporting systems. By standardizing, we can replicate easily technology between compliance programs, pooling resources to build a single set of data integrity checks. The facilities and laboratories that supply environmental compliance data will have clearer direction on what needs to be reported and can streamline their reporting systems. Commercial laboratories in our certification program have been receptive to the concepts in NR 148. Our regulated facilities are concerned about data quality, particularly if and when compliance data is posted on the Internet so even though reporting data quality may be an added burden initially, we anticipate there will be broad support for these proposals.

As the primary state agency responsible for environmental compliance programs, Wisconsin Department of Natural Resources (DNR) receives data for the following programs.

NPDES Permit	Underground Injection Control
Ambient Water Monitoring	RCRA (Hazardous Waste)
Storm water	Underground Storage Tanks
Non-point Sources	Brownfields
Groundwater Protection	Remediations
Safe Drinking Water	Air Emissions Inventory
Contaminated Sediments	Ambient Air Monitoring
Biosolids Management	Stationary Source Air Monitoring

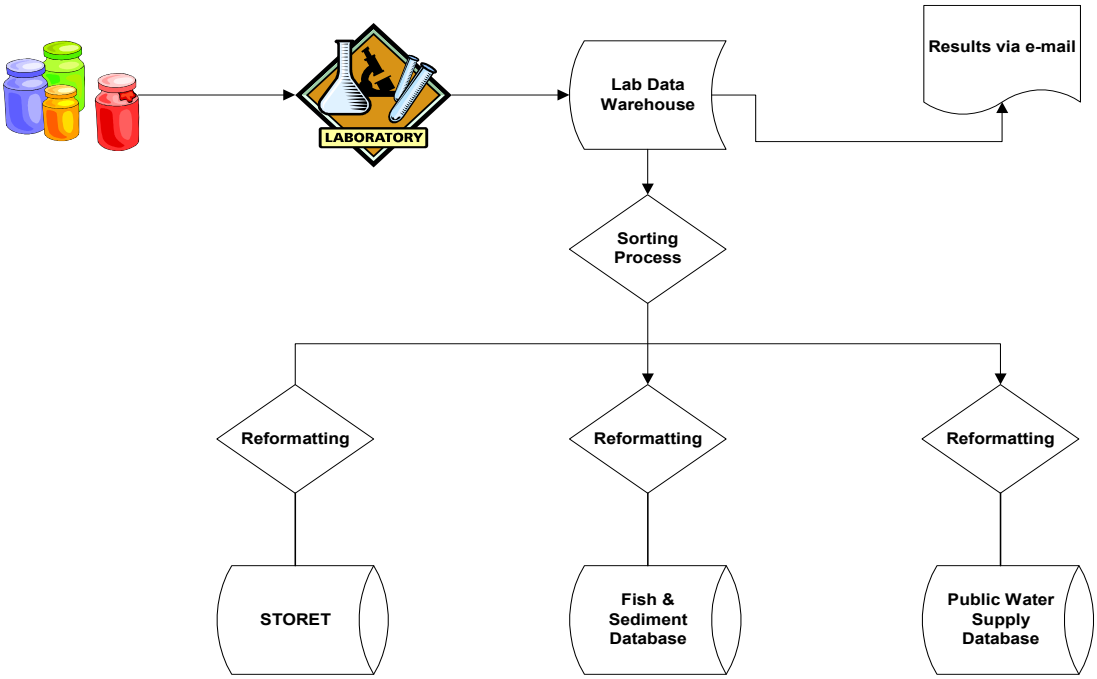
Given the large amount of data generated for these programs, DNR has designed databases to manage the information, to determine whether results exceed any permit limits and to verify that compliance schedules have been met.

During our Quality Management Plan development, we looked at the entire data collection and reporting system and what mechanisms ensured data quality. We used systems diagrams to identify the various steps in the process, who is involved, and their capabilities. Several of the programs involve self-monitoring for compliance and the level of sophistication ranges from wastewater treatment plant operators with high school education and small businesses to environmental consultants and multi-national corporations with dedicated regulatory compliance staff. In quality improvement terms, there are several supplier-customer relationships that affect the quality of the data.

High Tech

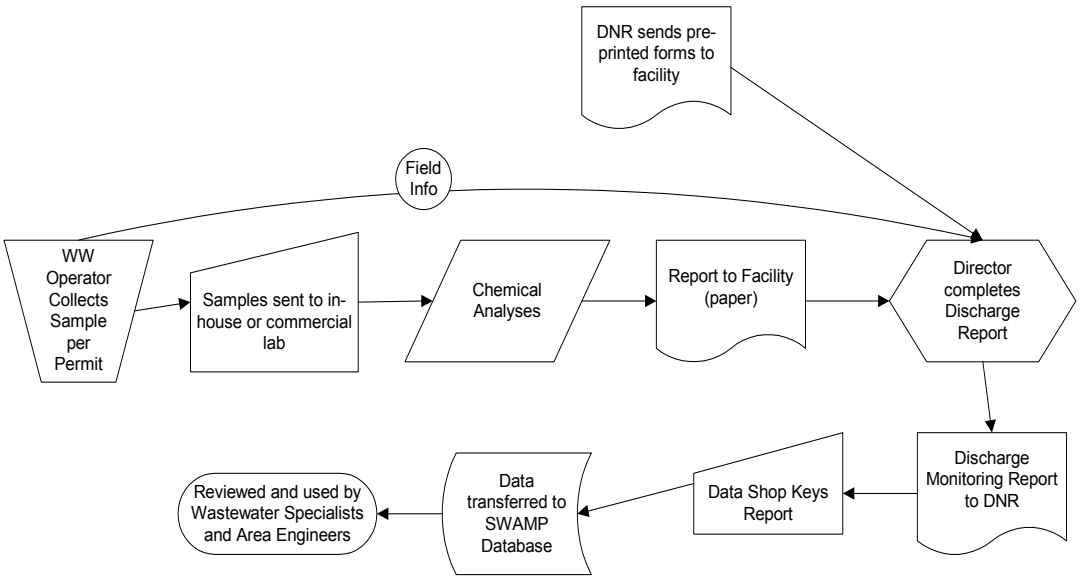
For hazardous waste and air emissions, facilities use a Consolidated Reporting System (CRS) to report their data to DNR. CRS is computer software that allows facilities to update their emissions information. For air, facilities supply production and process information related to their air emissions (e.g. natural gas usage for a boiler or aluminum processed through a furnace and % efficiency). For some facilities and processes, air emissions also may be based on stack testing results. For Hazardous Waste, the facility supplies their hazardous waste generator information and any transfer, storage or disposal quantities. The information on these reports is used to calculate environmental fees related to those emissions. Once the information is entered, the program includes a completeness check. Before December 2000, DNR provided facilities with the program and facility-specific information and the facilities sent diskettes with their reports back to DNR. In the December 2000 reporting year, DNR began taking greater advantage of the Internet. The software can be downloaded from our website and reports can be returned to DNR by mailing a diskette or by e-mail. Once received and uploaded into DNR's database, our program calculates annual air emissions based on the EPA emission factors

In one of our most direct but perhaps highest tech systems involving environmental samples, DNR staff collect samples, complete the transmittal form and deliver them to our State Laboratory. The State Lab staff log the samples into their LIMS and performs the requested analyses. As the analyses are completed and pass quality control review, results are transmitted electronically to DNRs lab data warehouse and staff are notified by e-mail that results are available through a Web browser. Once all sample results are complete the data is sorted and reformatted for the appropriate database.



Low Tech

In sharp contrast, the process for transmitting wastewater data as part of the National Pollutant Discharge Elimination System definitely fits the "low tech" mold. This system relies heavily on transcribing data from one paper report to the next before data is manually keyed and transmitted to DNR's database.



Approximately 170 of 1200 facilities generated paper DMRs onsite using various software packages. While the output looked similar to the DMRs we supplied, we still had to key the data manually and many had errors that caused the databases to reject the information. This resulted in a situation where 75% of the time needed to correct errors in the system was attributable to 14% of the forms. We asked facilities to stop using their electronically generated reports until we could devise a more rugged approach.

In-Between

Wisconsin's regulations for solid waste landfills require that groundwater monitoring results be reported electronically for facilities with more than four wells. Frequently, the landfill contracts with a consultant or full-service environmental laboratory to collect the samples and prepare the necessary reports to DNR. Like CRS, facilities send results to DNR on diskette with a paper certification, but the report is merely the required data in either a fixed width or comma delimited ASCII format. The data is uploaded to the database, checking for viruses in the process. The computer then determines whether there are problems with valid license numbers, sample locations (point IDs), and parameter codes and date relationships. If there are, the person doing the upload gets an error message and the data is retained in a "held" table until the issues are resolved.

Moving to the Future - "Let's Make a Deal"

As our budgets get tighter, we are under increasing pressure to reduce costs. Optimizing our data reporting systems seems like a prime candidate for streamlining. However, we face the ago old dilemma. How can we take advantage of current technology for data transfer in our environmental compliance programs if those supplying the data don't have computer access or have limited capabilities? At times, it seems like "Let's Make a Deal" on a giant scale. We see the future holding a number of choices for reporting to us.

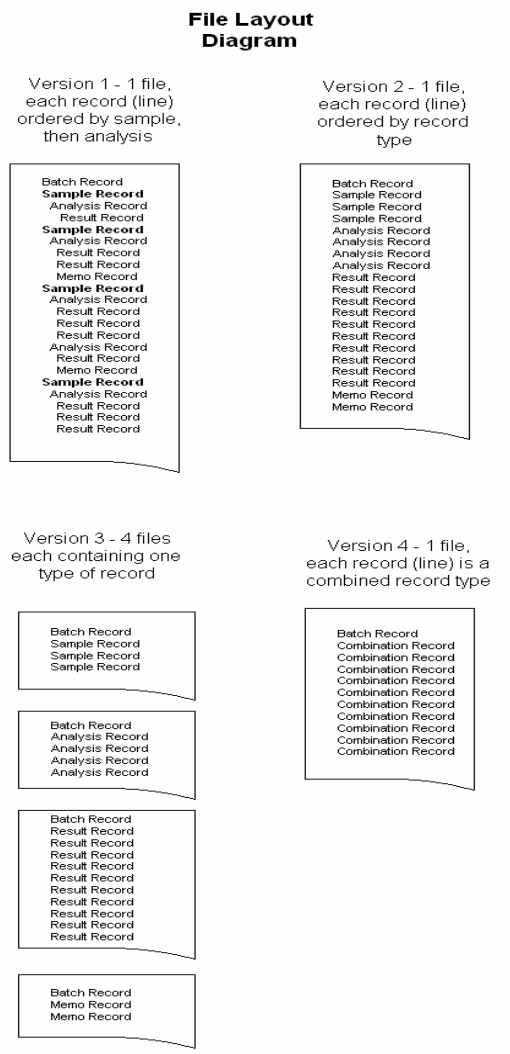
Behind Door Number One - e-DMR

In 1999, we surveyed facilities with permits. With almost two thirds of the facilities responding, 70% indicated that they had Internet access and 75% responded that they would choose Web forms over paper. We are beginning a piloting for an electronic Discharge Monitoring Report (DMR) that relies on completing web-based forms containing permit-specific information. The pilot will use a password/user identification with a secure socket link (SSL) security system. This will ensure the data transmitted to the Department is from an authorized agent of the permitted facility and that the data files cannot be intercepted or corrupted during entry on our server. The system assures data integrity through the use of a document key that is linked to the information when it is filed in the Department's System for Wastewater, Applications, Monitoring, and Permits (SWAMP) database. The document key is a mathematical total of key data points designed such that any changes in the data will impact the checksum total. When the data is sent to DNR, the facility prints a certification page that includes the key. The authorized person signs the certification and sends it to DNR who then scans the bar code, completing the process. In the next year, we hope to develop machine-to-machine data transfer using an Extensible Markup Language (XML) protocol for batch submissions.

Behind Door Number Two - Expanding the Lab Data Portal and Warehouse

The Safe Drinking Water program is testing procedures to use the Lab Data Portal for receiving compliance monitoring data directly from drinking water laboratories. We have defined the data structure

and outlined several formatting options, on using XML and four for ASCII tab-delimited formats. Although the diagram shows only analysis information, sample "header" information with facility and sample point data follows a similar formatting scheme and is linked by an appropriate key. Data sets are e-mailed as attachments to a specified mailbox, which sends an acknowledgment to the sender. A batch record identifies the facility and includes appropriate password information. Once the user's identity is verified, the data is transferred to the data warehouse.



We plane to develop a quality assurance module for the warehouse and expand data receipt capabilities to the other compliance programs.

Behind Door Number Three - Paper

At least of the near term, we don't see paper reports disappearing. We are concerned about the burdens placed on smaller communities and very small businesses. Some have suggested that DNR supply its surplused computers with reporting software to this sector to facilitate the transition to universal electronic reporting. That concept may one of our solutions in the future.

What's Next?

WI-MAP - Statewide Electronic Security

Wisconsin's Department of e-Government (DEG) is developing an e-Business security system for anyone that does business with the State. This system authenticates users and provides controlled access to State Government e-service. DEG evaluated commercially available authentication services and decided that it made fiscal sense to develop a system that would administered internally.

Standardized Reporting - Data Reporting Rule

DNR's current data reporting requirements are inconsistent and scattered across many Administrative Codes. Each regulatory program develops its own data systems and reporting conventions and as a result, regulated facilities that report to multiple compliance programs also need to develop multiple reporting systems. We propose to consolidate the analytical data reporting requirements into a single data reporting rule, NR 148, while changing the current rules to refer to NR 148. This would bring consistency to the Department's analytical data reporting and provide the regulatory framework for electronic submission of environmental monitoring data. The rule has four goals:

Consolidate compliance reporting requirements into a single Code
Standardize reporting conventions where possible
Specify the minimum required data elements, implement data standards
Identify performance criteria for electronic reporting

Presently, the individual programs receive and store only the information that they need for their immediate purpose. Present reporting systems provide sample results but do not include useful indicators of the quality or reliability of the data. Often, the data are inadequate to be applied appropriately to resource management decisions, other programs or compared to future data. There are clear economic advantages to locating and using appropriate existing monitoring data as opposed to generating entirely new monitoring data for a single purpose. In proposing a standardized set of information to be reported with sample results, the Department recognizes that although the goals of monitoring may be quite different, core information needed to evaluate that data is the same. By establishing data standards and bringing consistency to data reporting, the regulated community, analytical laboratories and the Department gain opportunities to streamline, increase data transferability, and minimize costs associated with reporting monitoring data. Financial and programming resources can be pooled to allow all parties to optimize their reporting systems.

Data Standards? - Definitely; Data Warehouses? - Maybe

DNR developed biological and chemical data standards as part of our Aquatic and Terrestrial Resource Inventory initiative. We are beginning to estimate the costs to implement those data standards throughout the Department. Several of us involved in the process see data warehouses as a cost-effective option for implementing the data standards, promoting data sharing between monitoring programs, providing a data layer for our GIS systems both internally and those available to the public, and for transferring data to EPA. Although it's a noble goal, the obstacles may prove insurmountable, at least for now. Sharing data will mean that we will need to institute change management procedures. Command and control are not popular management concepts these days. The various programs will need to give up some of the control over their data and spend time coordinating activities with people outside of their work area. In times when our resources are stretched thin and we are being held accountable for program-specific grant activities, our challenge will be to streamline processes with minimal disruption of our current activities. Our budgets may not allow for spending money now to save more money later.

The NEIC Document Control System

Kaye Mathews, Document Control/Records Management Coordinator
U.S. EPA National Enforcement Investigations

On February 1, 2001 the National Enforcement Investigations Center (NEIC) received accreditation for environmental measurements from the National Forensic Science Technology Center (NSFTC). This accreditation covered the specific areas of field measurements/monitoring testing, field sampling, and laboratory analysis. Included in these areas are such activities as evidence management, security, safety, fact and expert witness testimony, depositions, expert opinions, and affidavits. This accreditation also addresses document control/records management, procurement, facility environment; equipment inventory, maintenance, and calibration; and training, competency evaluation, and proficiency testing. This accreditation meets the requirements of ISO/IEC Guide 25 and ANSI/ASQC E4 and elements of the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) Manual.

Written documentation of how NEIC forensic and environmental measurement activities are conducted is central to achieving and maintaining accreditation and for implementing the NEIC Quality System. The collection of policies, operating procedures, guidance, and management plans provides a comprehensive guide to NEIC processes. An electronic distribution tool called DocLink provides desktop access to current and archived document versions, eliminating the need for hard copy distribution. DocLink also provides efficient processing, indexing, disseminating, and retiring capabilities.

Introduction

Effective February 1, 2001, the U.S. EPA - NEIC was granted accreditation for its overall environmental measurement activities that include: field sampling, field measurements and monitoring, and laboratory measurements. The accreditation criteria incorporate nationally and internationally accepted forensic and quality management standards.

The NEIC Accreditation Standard is based on ISO/IEC Guide 25 (the international guide for the competence of testing laboratories), ANSI/ASQC E4-1994 (the national standard adopted by U.S. EPA for quality management of environmental data collection), and references specific aspects of the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) Manual (a nationally recognized manual for organizations involved in forensic activities in the U.S.).

ISO/IEC Guide 25 contains a requirement for a documented quality system. This system must include procedures for indexing/numbering, dissemination, reviewing, revising, and retiring quality management documentation such as policies and procedures.

A pre-accreditation gap analysis audit conducted at NEIC in 1999 identified one significant deficiency: inadequate written quality management documentation. After this audit, NEIC formed teams to develop and/or revise policies, procedures, guidance, manuals, and management plans. The NEIC Quality Staff worked closely with the Records Management/ Document Control Coordinator and the NEIC Library's web development expert to develop an electronic desktop reference system instead of a hard copy distribution system for controlled documents.

Documents vs Records

ISO standards make a distinction between documents and records and NEIC has adopted these definitions for its document control practices. Controlled documents describe how work is expected to be conducted (i.e. policies and procedures). Records provide objective evidence of actions taken and observations made while implementing NEIC policies and procedures. NEIC has procedures for both document control and for records management. At NEIC, quality management, document control, and records management are all in the same organizational unit and they each report to the NEIC Director.

Document Development and Authorization

The first step in developing a document control system was to build a process for developing procedures. This process ensures a uniform approach to developing, reviewing, authorizing and publishing all documents related to the NEIC Quality System. Standardized templates, numbering, revisions, fonts, and file names also provide a uniform, consistent approach to individual documents.

Draft controlled document are reviewed prior to authorization by subject matter experts, affected parties, technical coordinators, and approving authorities. This review includes grammatical, editorial, technical, and legal content assessment. Controlled documents are authorized for implementation in writing by the appropriate approving authority within an organizational unit. A unique document control number is assigned to each new document for the life of the document and all revisions are designated R1, R2, etc.

Prior to the effective date of the document, briefings and/or training on the use of the document are required to take place. NEIC conducts training on each controlled document prior to implementation to assure that personnel are aware of the existence of the document and also aware of their implementation responsibilities.

Document Distribution System - DocLink

DocLink, the NEIC computer desktop distribution system, resides on the local area network drive, but web tools give it the look and versatility of an intranet site. Original electronic versions of documents in a variety of formats (WordPerfect, Corel Draw, PowerPoint) are converted to PDF format using Adobe Acrobat. The Homesite web authoring tool is used to create HTML pages which provide links to the PDF documents. The Netscape browser is used to pull up the HTML files.

The only currently authorized version of an NEIC controlled document is the Doclink version. When printed out, each document page is watermarked as a copy. This approach provides a central source for identifying the current version of a document, for preventing unauthorized changes to the official version, and for precluding the unintended use of obsolete documents.

DocLink is arranged by subject and also by document number. Hyperlinks are used in the document text to reference related documents. Large documents are book marked to facilitate navigation. Retired versions of controlled documents are archived in a specific section of DocLink for future reference. Key EPA and outside reference documents are also linked in DocLink.

Document Revision, Retirement, and Archiving

Controlled documents are periodically reviewed for accuracy and applicability, Records are kept of those reviews whether the documents are revised or not. Revisions to existing documents can be submitted in writing at any time by NEIC personnel to Quality Staff. Revisions to documents follow the same steps as original document authorizations through review and approval by approving authorities. A summary of revisions is indicated in the document revision history section.

A proposal to retire a document can be submitted in writing at any time by NEIC personnel to Quality Staff. Document retirement requests are reviewed and approved by the appropriate approving authority.

Retired documents reside in a specific section of DocLink and contain a watermark with “Retired” and the date.

Original signed controlled documents are maintained in a hardcopy master archive whether current, superseded, or obsolete. These documents are retained for legal and institutional knowledge preservation and are identified as current or archived versions.

Forms Management and Information Control

Forms A flexible document control process is used to provide commonly used NEIC-generated forms on DocLink. Formats and versions can easily be updated while still providing the most current version in a central electronic location.

Controlled Information Directives that provide information or direction which couldaffect the quality of the testing environment, or results, or who is in charge are managed as controlled information. Information related to the NEIC Quality System must be signed and dated by the author or appropriate person. Examples of controlled information are e-mails, memos, notices regarding equipment operation, sample storage, or access control, and the chain of command chart.

NV, Air & Waste Management Association, Pittsburgh, PA. Article available at: http://cluin.org/download/char/dataquality/qc_greatmyth.pdf

8. Popek, E.P. 1997. "Investigation versus Remediation: Perception and Reality" in Proceedings of WTQA'97—the 13th Annual Waste Testing and Quality Assurance Symposium, pp. 183-188. Paper available at http://cluin.org/products/dataquality/

9. U.S. Environmental Protection Agency (USEPA). 2000. OEI Quality System 2000: Office of Environmental Information Management System for Quality. http://www.epa.gov/oei/quality.htm

10. U.S. Environmental Protection Agency (USEPA). 2000. Guidance for Data Quality Assessment: Practical Methods for Data Analysis (QA/G-9 QA00 Update). EPA 600/R-96/084. July. http://www.epa.gov/quality/qs-docs/g9-final.pdf

Accreditation at the U.S. EPA - NEIC

Barbara A. Hughes, QA Coordinator, EPA/OECA/OCFT, National Enforcement Investigative Ctr.
K.Eric Nottingham, Laboratory Branch Chief and Jennifer Suggs, Chemist

The story of the NEIC accreditation process can be divided into four phases:

- *Thinking About It*
- *Planning It*
- *Achieving It*
- *Living It*

***Thinking About It** involved: asking ourselves if we really should seek accreditation and identifying a suitable accrediting body. **Planning It** involved: securing parent organization commitment, selecting a project leader and setting up a steering committee, establishing an internal communication mechanism, determining the requirements for accreditation, writing an accreditation standard, and conducting a gap analysis audit. **Achieving It** required: a lot of hard work including writing, reviewing, finalizing, and disseminating controlled documents; rewriting the quality policy and quality management plan; conducting training, competency evaluations, and proficiency testing; building a process to manage internal audits and corrective action requests; and conducting a pre-certification audit, a certification audit, and a verification audit. **Living It** includes: allocating resources, assuring the distribution of current quality management documents and information, conducting annual assessments of the quality system, and transitioning to more stringent accreditation requirements.*

Introduction

The U.S. Environmental Protection Agency - National Enforcement Investigations Center (NEIC) of Denver, Colorado is the specialty technical arm of the Office of Enforcement and Compliance Assurance (OECA) within the U.S. EPA. NEIC supports environmental enforcement through compliance assistance and civil and criminal investigations. OECA centralizes U.S. EPA environmental enforcement within one office. Within OECA, the Office of Criminal Enforcement, Forensics and Training is the headquarters unit for NEIC. NEIC is a center for technical support nationwide to state, local, tribal, and federal environmental enforcement and compliance assurance programs. NEIC is a source of expertise for technical analysis, compliance monitoring, engineering evaluations, forensic laboratory activities, information management, computer forensics, and courtroom testimony.

Effective 1 February 2001, NEIC was granted accreditation for its overall environmental measurement activities that include: field sampling, field measurements and monitoring, and laboratory measurements. NEIC became the first and only environmental forensic center in the U.S. to be granted this type of accreditation. This achievement required three years of intensive work by the entire staff to ensure that NEIC policies, procedures, guidance, manuals, and management plans were in accord with stringent accreditation criteria. The criteria incorporate nationally and internationally accepted forensic and quality management standards.

This pioneering accreditation, awarded by the National Forensic Science Technology Center (NFSTC) [1], is based on a standard that was cooperatively developed by NEIC and NFSTC. This rigorous standard was developed for a facility that conducts environmental measurements while adhering to forensic requirements in specific areas. The forensic areas in the standard include: evidence management, facility security; and testimony in trials, hearings, and depositions. This standard encompasses: document control and records management, procurement, safety, work place environment; equipment inventory, maintenance, and calibration; and training, competency evaluation, and proficiency testing. The NEIC Accreditation Standard is based on ISO/IEC Guide 25 (the international guide for the competence of testing laboratories), ANSI/ASQC E4-1994 (the national standard adopted by U.S. EPA for quality management of environmental data collection), and references specific aspects of the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) Manual (a nationally recognized manual for organizations involved in forensic activities in the U.S.).

NEIC sought accreditation because it would make a public statement about the quality of NEIC operations. Achieving accreditation demonstrates to our customers and to the general public that NEIC has taken a recognized and systematic approach to planning, conducting, documenting, and assessing forensic environmental data collection activities.

For the past 12 years, NEIC has been accredited for asbestos testing within the National Voluntary Laboratory Accreditation Program (NVLAP), operated by the National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce. The NVLAP accreditation is not discussed in this presentation.

NEIC Accreditation Process

The story of the NEIC accreditation process can be divided into four phases:

- Thinking About It
- Planning It
- Achieving It
- Living It

Thinking About It

In recent years, forensic laboratories have come under greater legal and technical scrutiny. There is a need to assure the legal community and the public that a measurement facility is capable of generating reliable and legally defensible data. NEIC decided that accreditation was necessary because it would make a public statement about the quality of NEIC operations. Achieving accreditation demonstrates to our customers and to the general public that NEIC has taken a recognized and systematic approach to planning, conducting, documenting, and assessing forensic environmental data collection activities. As we became more involved in the process, we came to understand that accreditation forces a facility to examine, understand, and continually improve its quality system.

Many accreditation bodies exist and there are variations in their focus that must be evaluated by the facility seeking accreditation. The NEIC accreditation for asbestos testing was insufficient in scope to encompass the broad spectrum of NEIC measurement activities. The NEIC was looking for an accreditation body that:

- Would accredit both field and laboratory operations
- Would accredit both forensic and environment measurement activities

- Could provide third-party independent audits
- Was covered by national and/or international requirements for accreditation bodies (e.g., ISO/IEC Guide 58 - *Calibration and Testing Laboratory Accreditation Systems: General Requirements for Operation and Recognition*)

There were accreditation bodies in the U.S. willing to accredit laboratory operations but not field operations. There were organizations that were prepared to accredit environmental measurements but could not address the forensic aspects of NEIC operations. Many organizations could conduct third-party audits but few were operating under national and/or international requirements for accreditation bodies. The NEIC eventually approached the National Forensic Science Technology Center and they agreed to customize an accreditation standard for our forensic environmental measurement activities that included both field and laboratory operations.

Planning It

The accreditation process is a long-term and labor intensive process. It requires a commitment by the parent organization to allow a facility to devote the funding and time needed for a successful outcome. NEIC was able to obtain this commitment from our headquarters office in located in Washington, DC. Once we had this commitment, NEIC selected a leader for the endeavor and set up a steering committee to:

- Provide direction and mobilize internal resources
- Develop and manage overall and task specific time lines to keep the process on schedule
- Assign tasks to staff
- Review assignment work products and resolve policy, procedure, or practice issues
- Arrange for external and internal assessments
- Formulate plans for responses to external and internal assessments

NEIC established a communication mechanism early in the process that included meetings of all personnel, regular electronic communications, and posters showing progress along time lines. The purpose of this mechanism was to keep personnel informed and involved and to show progress toward the eventual goal of achieving accreditation. NEIC made an effort to identify and talk regularly with staff who were less than enthusiastic about accreditation at NEIC. A conscious effort was made to give many of these persons assignments during the accreditation process in order to minimize their apprehensions about operating within a new system.

The NEIC accreditation time line included:

- Securing training on ISO/IEC Guide 25 and on how to conduct internal audits
- Developing the NEIC Accreditation Standard and conducting training on the standard
- Submitting to an initial gap analysis audit in order to define the work needed to bring written documentation and practices in line with accreditation requirements
- Developing tasks, teams, and time lines to address gap analysis audit findings

The initial gap analysis audit identified one significant deficiency: inadequate written quality management documentation. After this audit, NEIC formed teams to develop and/or revise policies, procedures, guidance, manuals, and management plans. We developed a document control system that provided an electronic desktop reference system instead of a hard copy distribution system. This approach enabled us

to have a system for identifying the current version of a document, for delivering that current version to the desktop computer of personnel, for preventing unauthorized changes to the official version, and for precluding the unintended use of obsolete documents.

Achieving It

During this phase of the process, NEIC focused on writing, reviewing, finalizing, disseminating, and implementing the controlled documents needed to define the NEIC Quality System under accreditation. The quality system was developed as a structured and documented management system that included: written policies, procedures, guidance, manuals, and management plans; management and staff authority, responsibility, and accountability; and an approach for generating and maintaining records that provide objective evidence of actions taken while implementing the system. NEIC conducted training on each controlled document prior to implementation to assure that personnel were aware of the existence of the document and also aware of their responsibilities in the implementation of the document.

At the heart of the quality system is the NEIC Quality Policy. This policy was written to reflect management philosophy on quality and stands as a guiding principle for NEIC measurement activities. The quality policy states that NEIC will operate with a quality system that supports its mission and goals. It also states that management is committed to supporting staff efforts in meeting customer expectations of quality.

The NEIC Quality Management Plan (or quality manual) is an essential component of the quality system. It was revised to describe the newly enhanced quality system. It describes the NEIC approach for quality management and provides a rationale for that approach and is therefore used to implement the quality policy. It is a general reference document that points to more detailed, task specific documents. The quality management plan is used as a reference when assessing whether the quality system is being successfully implemented.

NEIC identified the techniques that would be used in environmental data collection and defined training and competency evaluations for personnel in those areas. We set up internal and external proficiency testing programs in order to evaluate the continuing capability of personnel to participate in measurement activities.

Also during this phase of the accreditation process, NEIC implemented a formal process to conduct, record, and follow-up on internal audits on an annual basis instead of on an *ad hoc* basis. NEIC developed an annual audit schedule that covered assessing all aspects of the quality system. An electronic system was developed to track audits conducted, reports generated, auditors participating, observations that could lead to quality improvements, and findings requiring corrective actions. A process was developed to resolve any disputes resulting from these assessments, especially across organizational unit boundaries.

NEIC submitted to a pre-certification audit conducted by an external, independent team consisting of an ISO/IEC Guide 25 specialist, forensic specialists (county and state crime laboratory directors), and a U.S. EPA quality management specialist. The purpose of this audit was to determine what work still needed to be done to bring documents, practices, and records in line with accreditation requirements. This audit was our last chance to determine "accreditation readiness" before the certification audit. We developed plans and formed teams to address the pre-certification audit findings.

The certification audit team arrived at NEIC in December 2000. The team consisted of an ISO/IEC Guide 25 specialist, a forensic specialist from a crime laboratory, a U.S. Navy quality manager, and a U.S. EPA quality manager. The certification audit lasted for four days and included interviews,

observations, examination of facilities and equipment, and reviews of documents and records. The audit covered all organizational levels at NEIC including: Director/Deputy and Associates, Program Coordinators, Branch Chiefs, Project Leaders, Technical Staff, Administrative Staff, Quality Staff, and Safety/Waste Management Staff.

We developed plans to address the certification audit findings, which meant generating, assigning, and tracking additional corrective actions. The certification audit team leader returned to the NEIC in January 2001 to verify that all of the audit findings had been addressed satisfactorily, that corrective actions had indeed been implemented, and that corrective actions had not adversely affected other aspects of the quality system.

Success was achieved! The verification audit determined that NEIC was indeed in compliance with accreditation requirements. The audit team leader recommended to the NFSTC that NEIC be accredited.

Living It

NEIC is currently operating with a quality system that covers field and laboratory operations and forensic and environmental measurement activities. The quality system under accreditation addresses nationally and internationally recognized forensic and quality management requirements. This accreditation provides NEIC with the flexibility to develop or customize environmental data collection methodology appropriate to the variety and complexity of technical support needed to support its mission. Implementing the quality system under accreditation has led to several management and staff meetings to discuss specific work processes and the flow of work within those processes.

There are continuing costs associated with accreditation for: fees, training, technical assistance, external assessments, equipment replacement, facility improvements, proficiency testing materials, and management information system enhancement. Some staff resources must be dedicated to overseeing the implementation of the quality system and some staff resources must be expended to assess the system annually.

NEIC accepts that accreditation is not a substitute for "sound science", but it does provide a functioning system for generating "sound science". Accreditation of the quality system provides a system of internal and external checks on the processes used to generate scientific work products. NEIC understands that accreditation is only as good as our implementation, and that we are still required to defend the quality of individual work products associated with environmental enforcement actions.

The next step for NEIC in accreditation is the transition to an accreditation and a quality system based on ISO/IEC 17025 (the international standard for the competence of testing laboratories) and applicable international guidelines for forensic testing facilities by 2003.

References

- *ISO/IEC GUIDE 25: General requirements for the competence of calibration and testing laboratories*, 3rd ed., International Organization for Standardization, 1990.
- *American National Standard ANSI/ASQC E4-1994: Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*, American Society for Quality Control, Energy and Environmental Quality Division, Environmental Issues Group, 1995.
- *Laboratory Accreditation Board Manual*, American Society of Crime Laboratory Directors, 1994.

---

1. The mention of any organization, including but not limited to the National Forensic Science Technology Center (NFSTC) does not constitute an endorsement by the U.S. EPA. NFSTC Home Page, <http://www.nfstc.org/>(November 26, 2001).

Improved Quality Data Systems through the use of Standard Electronic Data Deliverables (EDDs) and Environmental Data Assessment Software

Pamela A. Wehrmann, Senior District Chemist, U.S. Army Corps of Engineers, and
Richard M. Amano, Principal Chemist, Laboratory Data Consultants

One of the challenges facing professionals in the environmental arena today is the collection and assessment of large amounts of environmental analytical data. The assessment of the quality of that data is essential as multi-million dollar decisions for project are made based on the analytical results. Also critical to environmental programs is the sharing and access of data across multiple data users. Standardization of electronic deliverables allows for collection of data from multiple data collectors into a single database for use by numerous data users or stakeholders on a project. This presentation will discuss the benefits of using a single EDD deliverable format and use of environmental data assessment software tools to do project planning and data assessment throughout the duration of the environmental project.

This presentation is a brief overview of the improved chemical data management available for environmental projects through use of a standard Electronic Data Deliverable (EDD) specification developed for the Army Corps of Engineers, Sacramento District and support software modules that allow for review and assessment of the analytical data collected for a project.

The EDD provides a standardized format for the collection of chemistry data. This format allows for two benefits. The first benefit is found in streamlining at the laboratories to produce data deliverables that can be verified immediately for completeness and compliance against project specific data quality indicator criteria as defined in the project specific Quality Assurance Project Plan (QAPP). The second benefit in use of a standard format allows for multiple contractors to work on a project over the life of the project and all submit and share the environmental data in a single compatible data base. The a standard format deliverable allows a comprehensive data base for the project, or multiple projects, regardless of which contractor collected the data, which lab analyzed the data, and which phase of the process the project is in. Data collected from portable field labs and from multiple fixed site labs can be combined into a single database that can be queried and from which the analytical data can be managed and evaluated for trends.

The support software consists of the Contract Compliance Screening (CCS), Automated Data Review (ADR), and Environmental Database Management System (EDMS) software programs developed by Laboratory Data Consultants, Inc. under contract to the Army Corps of Engineers, Sacramento District. The software programs use an electronic data deliverable (EDD) format based upon data elements originally documented in the Implementation Guide for the Department of Energy Environmental Management Electronic Data Deliverable Master Specification (DEEMS). The software has been recently updated to accept an EDD deliverable in the new Superfund Electronic Data Deliverable (SEDD) format. The software was developed on a Microsoft ACCESS 97 platform and has been updated to ACCESS 2000. Customized modules perform automated data review equivalent to an EPA Level 3 validation and provide the user with discrete data qualification flags. The qualified data is exported into a master database for overall project use.

Once Technical Project Planning is complete and Data Quality Objectives (DQOs) and Data Quality Indicator (DQI) requirements have been determined for the project, those requirements can be documented for the Quality Assurance Project Plan (QAPP) by setting up project libraries in the ADR module. The project library is then sent to the lab for the labs use in screening the EDDs for contract compliance errors against the requirements set up for the project in the QAPP. The EDD format includes QA/QC batch links and routine accuracy and precision parameters such as surrogate, matrix spike, and laboratory control sample recoveries. In addition, initial and continuing calibration and GC/MS tuning data can be delivered in this format. Development of the EDD integrated these data elements required by end users with consideration for the current data deliverable capabilities of commercial laboratories. Once screened, the EDD deficiencies are detailed in an outlier report. Access to the EDD file in table format allows for quick and easy correction of errors. When corrected and error free, the EDD is sent to the client for automated data validation/review with ADR. This saves project time and money by assuring that an EDD comes from the lab that is as error free as possible.

The ADR software is initiated by the data user (i.e., Army Corps chemist, prime contractor chemist, etc.) to review analytical data based upon the project specific QAPP criteria. Upon execution of the program, data is qualified using standard EPA data flags and exported into a master database. Command buttons generate reports such as a rejected data table, method blank contamination, surrogate outliers, etc. Forms and view screens also provided on-line review of data qualifiers. Automated data review processes save projects time and money by allowing all of the data to be reviewed and not just a "representative" portion of the data set. This is truly cost effective on large projects where review of mountains of hard copy data can be a daunting, if not impossible, task. The Sacramento District has seen an almost 50% reduction in validation costs by automating data validation on large ground water monitoring projects. Automated data review also allows for nearly real-time review of analytical data quality issues so that data gaps can be assessed and addressed quickly.

The Environmental Database Management System (EDMS) compiles the validated data downloaded from the ADR system. The database program has user functions which allows for comparison of primary data versus QA split lab data, comparison of results against project action limits or Preliminary Remediation Goals (PRGs) or Maximum Contaminate Levels (MCLs), and calculates the completeness values for each test over any period of time. The four types of completeness values include contract, analytical, technical, and field sampling completeness.

In summary, the CCS, ADR, and EDMS software programs were developed as tools to support technical staff in the validation and evaluation of analytical chemistry data using an expedited and cost effective automated process. The EDD provides a standardized format for all of the project data and allows for data from various contractors to be combined into a single database. The standard EDD format allows for streamlining at the laboratories to produce data deliverables that can be verified immediately using CCS software for completeness and compliance against project specific data quality criteria, and errors can be immediately corrected. The EDMS allows the data end user to efficiently evaluate large data sets for key indicators and ultimately determine the usability of the data for making project decisions.

Required Steps for the Validation of a Laboratory Information Management System

Elizabeth Turner, Laboratory Chief
Jojean Bolton, Quality Control Officer
USACE Washington Aqueduct

The task of managing laboratory data is not a new one. Over the past two decades, the use of Laboratory Information Management Systems (LIMS) has revolutionized how laboratories manage their data. A LIMS is more than software; it has become the workhorse of the laboratory encompassing laboratory workflow combined with user input, data collection, instrument integration, data analysis, user notification, and delivery of information and reporting. Types of organizations that utilize LIMS vary greatly from research laboratories to manufacturing laboratories to environmental testing laboratories.

Commercially-available LIMS have been around since the 1980's. In addition, many laboratories have designed, implemented, and maintained in-house LIMS. The heart of any LIMS is the software. Like other laboratory systems, the LIMS software is subject to quality control and quality assurance checks. In regulatory environments, this associated QA/QC is referred to as "system validation." The primary purpose of system validation is to ensure that the software is performing in a manner for which it was designed. For example, the system acceptance criteria should be established and tested against quantifiable tasks to determine if the desired outcome has been achieved. LIMS features, such as autoreporting, reproducibility, throughput, and accuracy, must be quantifiable and verifiable. System validation ensures that the entire system has been properly tested, incorporates required controls, and maintains and will continue to maintain data integrity. Laboratories must establish protocols and standards for the validation process and associated documentation. Although vendors of commercial LIMS perform initial internal system validations, the system must be revalidated whenever the end user, vendor or third party adds modifications or customizations to the LIMS.

Currently, detailed guidance regarding system validation of LIMS is not available to the user. The issue is addressed in Good Automated Laboratory Practices (GALP) and National Environmental Laboratory Accreditation Conference (NELAC) documents which indicate specific requirements or recommendations for operational checks and periodic testing, however, it is up to the laboratory to determine suitable methods to accomplish these tasks. Proper validation of a LIMS will allow a laboratory to comply with regulations and also provide comprehensive documentation on the system that is necessary to troubleshoot future problems.

The validation of a Laboratory Information Management System (LIMS) is becoming an increasingly important issue for many laboratories. Over the past several years, guidance and regulations addressing data system validation have been developed and applied. For the most part, the guidance documents such Good Automated Lab Practices (GALP) and USFDA Good Manufacturing Practices, 21 CFR part 11

(Electronic Signatures) and Quality System apply primarily to the food and pharmaceutical industries. In August 2001, the US Environmental Protection Agency (USEPA) proposed the Cross Media Reporting and Record-Keeping Rule for USEPA environmental programs and laboratories. Validation is a pre-requisite for all systems that handle regulatory data and is a fundamental requirement for compliance with 21 CFR part 11 and similar protocols. The objective of the validation process is to ensure that a system does what it purports to do and will continue to do so. Validation not only satisfies regulatory requirements but is also an excellent tool for organizations to use so that they can be confident that the LIMS performs the way it is expected to perform.

There is no single, standard way to plan and implement a validation process. The one universal truth in the validation process is that validation activities need to be conducted throughout the entire LIMS life-cycle. The validation process starts with the functional requirements development phase in the purchase of a LIMS and continues through specification, testing, implementation, operation and retirement of a system.

Principles of Software Validation

There are ten general validation principles that are applicable to a LIMS.

1. *Timing:* Validation is not a one time event. It should begin when planning and input for a system begin. Validation does not end until the product is no longer used.
2. *Management:* Proper validation of a LIMS includes the planning, execution, analysis and documentation of appropriate validation activities and tasks.
3. Plans: Established design and development plans should include a specific plan for how the software validation process will be controlled and executed.
4. *Procedures:* Validation procedures should be developed and documented. The validation process should be conducted according to the established procedures.
5. *Requirements:* To validate a LIMS there must be predetermined and documented requirements. If a request for proposal (RFP) was thoroughly developed, it will contain the requirements necessary for validation.
6. *Testing:* Verification includes static and dynamic techniques. Static techniques include paper / document reviews, dynamic techniques include physical testing - demonstrating the system's run time behavior in response to selected inputs and conditions. Dynamic analysis alone may not provide sufficient information to determine if the system is fully functional and free of avoidable defects. Static techniques are used to offset limitations of dynamic analysis. Inspections, analyses, walk-throughs and design reviews may be more effective in finding, correcting and preventing problems at an earlier stage of the development process.
7. *Partial validation:* A system cannot be partially validated. When a change is made to the system, the validation status of the entire system should be evaluated.
8. *Amount of effort:* The magnitude of the validation effort should be commensurate with the risk associated with dependence on critical function. The larger the project and staff involved, the greater the need for formal communication, more extensive written procedures and management control of the process.
9. *Independence:* Where possible, the validation activities should be conducted using the basic quality assurance concept of "independence of review".
10. *Real world:* It is fully recognized that LIMS are designed, developed, validated and regulated in a real world environment. Environments and risks cover a wide spectrum and each time a validation principle is used, the implementation may be different.

Life Cycle Activities

Activities in a typical LIMS life-cycle include:

Management / Project Initiation Phase
Requirements Phase
Design Phase
Implementation Phase
Installation and Test Phase
Operation and Support Phase

The process of validation needs to start at the beginning of the LIMS lifecycle. Performing validation at the end of the process would add three to five months to a project and would preclude use of the LIMS during validation. For each of the life-cycle phases, certain validation tasks are performed.

Management / Project Initiation Phase

During design and development planning, a validation plan is developed to identify required validation tasks, procedures for reporting anomalies and their resolution, and resources needed for validation and management review requirements. The validation plan should include:

Specific validation tasks for each life-cycle activity
Methods and procedures for each validation task
Acceptance criteria for each validation task
Inputs for each validation task
Outputs from each validation task
Criteria for defining and documenting outputs
Roles, resources and responsibilities for each validation task
Risks and assumptions

21 CFR 820 requires that management identify and provide the appropriate validation environment and resources. Each validation task will require personnel as well as physical resources. The validation plan should identify the personnel, facility and equipment resources for each validation task. Procedures should be created for the review and approval of validation results including the responsible organizational elements for such reviews and approvals.

Requirements Definition Phase

A LIMS requirement specification document should be created with a written definition of the software functions to be performed. It is not possible to validate a LIMS without predetermined and documented requirements. Typical requirements specify the following:

All inputs the system will receive such as sample collection data

All outputs the system will produce such as reports

All functions the system will perform such as calculations and trend analysis

All performance requirements that the system will meet (data throughput, reliability, timing, etc.)

The definition of all internal, external and user interfaces (instruments, other software)

What constitutes errors and how errors should be handled

The internal operating environment for the software (hardware platform, operating system, etc.)

All ranges, limits, defaults and specific values the software will accept

A software requirements interface analysis should be conducted, comparing the software requirements to hardware, user, operator and software interface requirements for accuracy, completeness, consistency, correctness and clarity to ensure that there are no inconsistencies. During the requirements definition phase, validation activities should ensure that the requirements are testable. A thorough request for proposal could serve as your requirement specification document.

Design Phase

The design phase of the LIMS life-cycle is performed by the LIMS vendor. In order to ensure that the vendor performs the required validation activities, the laboratory may request a vendor audit and specify the validation requirements in the request for proposal. The primary goal of the audit is to ensure that the vendor's software development and management procedures are consistent with accepted practices and are traceable to a reference point. Often a vendor will be certified by an independent auditor as complying with ISO 9000 standards. A copy of the LIMS vendor's ISO 9000 certificate may be sufficient to meet the design phase validation requirement.

Implementation Phase

Implementation is the activity where detailed design specifications are implemented as source code. For commercial LIMS products, the implementation phase is carried out by the vendors' programming departments.

Installation and Test Phase

Installation testing is an essential part of validation for a LIMS. Terms such as beta test, site validation, user acceptance test and installation verification have all been used to describe installation testing. Installation testing is any testing that is conducted at a user's site with the actual hardware and other software that will be part of the installed LIMS configuration. The testing is accomplished through either actual or simulated use of the software being tested within the environment in which it is intended to function. If the computers for the LIMS are placed on a network, a validation procedure for proper network installation and connection will need to be established. All modems, printers, fax machines and instruments for data acquisition will need to be tested for proper integration.

Test plans should be created prior to the system installation. They should identify the test schedules, test environments, resources (people, tools, etc.), methodologies, cases (inputs, procedures, outputs, expected results), documentation and reporting criteria. Individual test cases should be associative with particular specification elements and each test case should include a predetermined, explicit and measurable expected result derived from specification documents in order to identify objective success / failure criteria. Installation testing should follow the predefined plan with a formal summary of testing and a record of formal acceptance. There should be retention of documented evidence of all testing procedures, test input data and test results. There should be evidence that hardware and software are installed and configured as specified. The testing phase should continue for a sufficient amount of time to allow the system to encounter a wide spectrum of conditions and events in an effort to detect any latent faults which are not apparent during normal activities.

The number of test plans will depend on the complexity of the LIMS and the level of detail required to adequately test the key features. Each externally visible function and each internal function should be tested at least once. Detailed written procedures and checklists are often used to facilitate consistent

application of intended testing activities. The methodologies used to identify test cases should allow for a thorough examination of the LIMS application.

In addition to an evaluation of the ability of the LIMS to properly perform its intended functions, there should be an evaluation of the user's ability to interface with it. Users should be able to perform the intended operations and respond in an appropriate and timely manner to all alarms, warnings, errors, etc.

There are several areas that require special attention during a LIMS validation study. All procedures for data entry and receipt of data, either electronic or manual, must be formal and implemented to ensure consistent execution. Data ranges should be developed as well as system edits activated to limit the introduction of erroneous data. Each input field should be identified and the allowable input defined. Testing should be conducted to demonstrate accurate processing of valid or "acceptable" data and the rejection of invalid or "unacceptable" data. If the LIMS performs calculations, evidence of the reliability of the formula / algorithm must be documented. Documentation often takes the form of published algorithms. Most hard coded calculations are preexisting formulae and can be traced to a published source.

All reports generated by a LIMS must conform to be established criteria for format, content and accuracy. Pre-established criteria may be part of a LIMS RFP. The specification document for each system output should describe the contents and format of the report. For a user-defined report, the report details should be defined and configuration management documented.

Operation, Maintenance and Support Phase

After a system is validated and becomes operational, changes, which may impact validation status, will occur during its operational lifetime. Any change to the LIMS should trigger consideration of revalidation of the system. There may be instances where no revalidation would be necessary after a change. One way to evaluate the need for revalidation is to review the impact the change would make to data accuracy, security and integrity.

Examples of changes to a system include hardware maintenance and upgrade, upgrade of the operating system and evolution of the LIMS application overtime. Configuration management is a set of procedures to insure adequate identification, control, visibility and security of any changes made to hardware, firmware, network, program source code or any specialized equipment associated with the LIMS.

The process of configuration management is quite simple. The initial system configuration is thoroughly documented. A listing of all components of the system should be compiled: includes all the release numbers, serial numbers of the application software and the operating system. The components comprising the hardware such as disks, memory, type of CPU and any peripherals as well as any documentation that is used with the system should be listed as well. As modifications to the system configuration are made the information should be recorded in the configuration log.

Change control defines responsibilities and documents the process of change. Change can include the resolution of system bugs and errors. The impact of the change to the LIMS needs to be evaluated. Items to consider include: time required to make the change, cost of the change, resources to make the change and benefits of the change. The effects of the change should be documented. Operational logs need to be updated to reflect any changes.

Validation activities are conducted throughout the entire LIMS lifespan. It starts with the requirements phase and continues through the operational phase. Even when a system is retired, it must be noted in the operational logs for future reference. Proper validation of a LIMS system will allow a laboratory to

comply with regulations and also provide thorough documentation on the system that will be needed to help troubleshoot the system.

References

ASTM. Guide E2066-00 Standard Guide for Validation of Laboratory Information Management Systems. 2001

EPA. Good Automated Laboratory Practices, Recommendations for Ensuring Data Integrity in Automated Laboratory Operations. Washington, D.C. August, 1995.

FDA. Quality System Regulation. 21 Code of Federal Regulations part 820, 1996.

McDowell, RD. Operational Measures to Ensure Continued Validation of Computerized Systems in Regulated or Accredited Laboratories. Lab Automat Info Manage 31, 1995.

National Institute for Science and Technology. Software Verification and Validation: Its Role in Computer Assurance and Its Relationship with Software Project Management Standards. 1996.

Paszko, C. and Turner, E. Laboratory Information Management Systems, Second Edition, Revised and Expanded. Marcel Dekker. 2001.

Comparison of EPA's QMS to SEI's CMMI®

Paul Mills, Leslie Braun and Don Marohl, DynCorp

EPA and other government organizations make decisions based on environmental measurements. How good are the data? How well are the data generators performing? What measurements apply to them? How can the data lifecycle processes be improved so data generators can continually provide the best data?

EPA's Quality Management System requirements go beyond evaluation of environmental data quality itself to examine systems associated with production, collection, processing (validation/verification), transfer, reduction, storage, and retrieval of data throughout a lifecycle. This QMS specifies minimum quality requirements for particular environmental programs. But how can you measure and compare programs that go well beyond the minimum, towards optimal quality?

This paper compares EPA's requirements for Quality Management Systems (R2) and Project Plans (R5) to the Software Engineering Institute Capability Maturity Model (CMM®). The CMM/CMMI® model provides for growth (staged or continuous) and a comprehensive assessment that is not yet provided in EPA's R2 or R5. Properly implemented, the CMMI® model serves as a quality framework for integrating and aligning organizational processes and implementing a program of continual process improvements. It identifies process areas ("things to do"), and provides measures of performance ("how well things are done") against specific goals and practices.

CMMI® uses a Systems Engineering Management approach, built on Process models, that helps identify "how good" the system is. Goodness is defined as stages in a complete model for optimal operation. CMMI® provides two methods for evaluating the goodness of the project. The Staged model in CMMI® provides a Maturity Level that is a well-defined evolutionary plateau describing the manner in which a specified set of processes are performed. As the organization advances in maturity, these levels become more defined and processes are tailored for specific project needs. The other method is called the Continuous Model in CMMI®, and it allows you to achieve Capability Levels. These are used to describe how well each project is doing in relationship to the different process areas. There are 6 Capability Levels from 0-5 that apply to individual process areas. Organizations using the Capability Level approach can select individual process areas that are important to specific projects, and work to improve the processes. Improving capability in individual process areas raises the organization's overall quality of products delivered. The Continuous Model, unlike the Staged Model, lets you pick higher maturity level process areas before completing all of the ones below.

Environmental measurement programs need to focus on the quality of the systems where data are collected, processed, transferred, etc. DynCorp built on the quality foundation from our experience with R2 to successfully implement CMM® practices in the development of Forms II Lite and other applications. DynCorp is now migrating to the CMMI® model that has evolved from the existing CMM® model. The CMMI® model focuses on the full cycle of Requirements Management from collection, refinement, analysis and validation throughout a project life cycle. It also has a more refined focus on the collection, analysis, and evaluation of meaningful measurements, so the results can be used to improve a process or product.

INTRODUCTION— What Are Capability Maturity Models?

Processes produce products. Process improvement is based on the premise that product quality is highly dependent on the processes used in its development. Capability Maturity Models are organized collections of best practices, based on the work of Crosby, Deming, Juran, and Humphrey. As a means to measure organizational maturity, these models present a systematic ordered approach to process improvement, and have demonstrated significant returns on investment in productivity and quality. Maturity models provide

Best practices
Measurement standards
Improvement paths

The Capability Maturity Model Integrated (CMMI®) describes the principles and practices underlying systems and software process maturity, and is intended to help software organizations improve the maturity or capability of their software processes that follow an evolutionary path from ad hoc, chaotic processes to mature, disciplined software processes. The CMMI® model has process areas with goals, and specific practices to meet these goals, and sub-practices that further define what is needed.

CMMI® identifies requirements for processes: what should be done, not how to do it. This is where the tailoring is applied for each unique project. The goals and specific practices are based on those whose effectiveness has been demonstrated. Appraisal methods based on a CMMI® allow the maturity or capability level of a project to be assessed, addressing each process area.

Properly implemented, CMMI® can provide the quality frameworks for integrating and aligning organizational processes, and implementing programs of continual process improvements. It aids in identifying process areas ("things to do"), and provides measures of performance ("how well things are done") against specific goals and practices.

- *Capability levels* are groups of practices that work together to enhance the capability of any process
- *Maturity levels* are groups of processes that together contribute to a step of improvement

Levels have generic goals and practices—"How well things are done." Process areas have specific goals and practices—"Things to do"

HOW DOES IT RELATE TO EPA WORK?

CMMI® will assist in raising the efficiency and effectiveness of processes to produce desired quality outputs. It is a way to break down each process to individual elements that can be managed and measured to enhance performance. EPA's Quality Management System requirements go beyond just environmental data to examine systems associated with data production, collection, processing (validation/verification), transfer, reduction, archival, and retrieval, throughout a lifecycle. This QMS could be viewed as specifying "minimum quality requirements" for particular environmental programs. In comparing the R2/R5 model to CMMI®, the following observations were made. Management and Organization in EPA R2 identifies the general objectives and goals of a quality system. It includes the scope, applicability and management responsibilities of the organization's quality system. It maps to both the CMMI® Project Planning and Project Monitoring and Control and the Process and Product Quality Assurance process areas. CMMI® Project Planning and Project Monitoring and Control organizations plans are created at the project level and include all the planning parameters such as staffing, directing, budgeting, coordinating and reporting. CMMI® Process and Product Quality Assurance it provides the framework for structuring the QA system such that staff and management are provided with objective insight into the processes and associated work

products. Both EPA and CMMI® include a process for identifying, reporting and tracking a non-compliance issue to closure.

WHAT ARE ITS ADVANTAGES?

The CMMI® identifies the process areas or "things to do" by defining specific and generic goals of the process and the sub-practices that can be used to support you in achieving the goal. EPA defines them as specifications as to what components the quality system should contain. The Maturity Level defines how good the system is. As your process matures you go from having repeatable processes in place for a few projects to having processes in place across the organization which allows you to do things more efficiently, maintain better control of your projects, provide better estimates, and improve the quality of your processes and products. The model is flexible in that it allows you to select a staged or continuous approach depending on your companies need. With the staged approach as you progress you advance to the next level. With the continuous approach you can identify the critical processes within each level and focus on the specific ones that are important to your success. EPA R2 andCMMI® both address the need to identify improvement activities. CMMI® addresses this by integrating improvement into all areas of the model through the use of measurement data, not just as an output of assessment, where R2 places it. CMMI® uses measurement as a critical tool in developing a measurement plan. Through a planning and analysis process, you define what you will measure, what measurements to take, how to collect them, who will collect them and the reporting frequency and analysis of metrics. Measurements are taken throughout the process life cycle (planning, development, testing and implementation). You analyze the results and use them to improve the process/product.

DynCorp built on the quality foundation from experience with R2 and was able to quickly identify the differences between R2 and CMMI®. This allowed us to prioritize what needed to be done and direct our resources to successfully implement CMM® practices in the development of Form II Lite and other software applications.

HOW DO I MEASURE CMMI® EFFECTIVENESS?

Project measurements must address common issues of schedule and progress; resources and cost; product size and stability; product quality; process performance; technology effectiveness; and customer satisfaction. An effective measurement approach—

Helps project and technical managers anticipate what can go wrong
Objectively supports tradeoff decisions if things go wrong
Helps to evaluate and communicate actual performance results
Ties in Quality Costs if right metrics are chosen
Uses Management Information Systems and Reporting
Embodies the Design part of EMS

In December, 2000, the CMMI Project released the Capability Maturity Model ® -Integrated (CMMI^{SM}) for Systems Engineering/Software Engineering, Version 1.02. One new feature contained in this model is the Measurement and Analysis Process Area. The practices associated with the first of three goals, "Align Measurement and Analysis Activities," establish the plan for measurement and analysis. They address: Why are we measuring (Establish Measurement Objectives)? What are we going to measure (Specify Measures)? How are we going to measure (Specify Data Collection and Storage Procedures)? And, what will be done with the data once we have it (Specify Analysis Procedures)? The practices associated with the second goal, "Provide Measurement Results," will result in getting the results of performing measurement and analysis into the hands of those who will take action and make decisions. Practices include: Collecting, Analyzing, Storing, and Communicating Measurement Data and Results. Activities associated with the third goal, "Institutionalize a Managed Process," include:

- Establish an Organizational Policy
- Plan the Process
- Provide Resources
- Assign Responsibility
- Train People
- Manage Configurations
- Identify and Involve Relevant Stakeholders
- Monitor and Control the Process
- Objectively Evaluate Adherence
- Review Status with Higher-Level Management

To begin implementing the CMMI[SM] for environmental programs:

- Review and document your existing measurement and analysis activities and procedures
- Evaluate the value of your existing measurement and analysis activities
- Integrate Measurement and Analysis into your processes**,** assign responsibility and train
- Establish an organizational infrastructure to support measurement and analysis

Capability Maturity Levels Compared to the 10 Elements in EPA's QMP, and People-CMM Elements

EPA QMS		Organizational Maturity Level			PEOPLE-CMM
Section	Element	Level	Characteristic	Key Process Areas	Key Process Areas
2	Management and Organization	1	Initial		
3	Quality System Components	2	Repeatable	Requirements Management Project Planning, Tracking, and Oversight Subcontract Management QA and Configuration Management	Compensation Training Performance Management Staffing Communication Work Environment
4	Personnel Qualification and Training	3	Defined	Organization Process Focus, Organization Process Definition, Training Program, Integrated Software Management, Software Product Engineering, Intergroup Coordination, and Peer Reviews	Participatory Culture Competency-Based Practices Career Development Competency Development Workforce Planning Knowledge and Skills Analysis
5	Procurement of Items and Services	4	Managed	Quantitative Process Management, Software Quality Management	Organizational Performance Alignment Organizational Competency Management Team-Based Practices Team Building Mentoring
6	Documents and Records	5	Optimizing	Defect Prevention Technology Change Management	Continuous Workforce Innovation Coaching Personal Competency

7	Computer Hardware & Software	2	Repeatable	Software Product Engineering	
8	Planning	2	Repeatable	Software Project Planning, Software Project Tracking and Oversight	
9	Implementation of Work Processes	2	Repeatable	Software QA	
10	Assessment and Response				
11	Quality Improvement	2	Repeatable	Software QA	

References:

CMMI Product Development Team, *CMMI SM for Systems Engineering/Software Engineering, Version 1.1 (CMMI-SE/SW, V1.02) Staged Representatio*n, Carnegie Mellon University, CMU/SEI-2000-TR-031 or ESC-TR-2000-096, December, 2001
ISO/IEC15939, *Information Technology – Software Measurement Proces*s, Committee Draft, December 2000.
Mark C. Paulk, Charles V. Weber, Bill Curtis, and Mary Beth Chrissis, *Capability Maturity Model: Guidelines for Improving the Software Process.* Carnegie Mellon University, Software Engineering Institute, 1995
SM CMMI and CMM Integration are service marks of Carnegie Mellon University.
® CMM and Capability Maturity Model are registered in the U.S. Patent and Trademark Office.
U.S.EPA 2001, *EPA Requirements for Quality Management Plans* (QA/R-2), EPA/240/B-01/002, March 2001
U.S.EPA 1999, *EPA Requirements for Quality Assurance Project Plans* (QA/R-5), EPA/600/R-99-033
"The Measurement and Analysis Process Area in CMMI" by Dave Zubrow

A Menu of Quality Systems: From TV Dinners to Filet Mignon

William Telliard and Marion Kelly, U.S. EPA Office of Water
Louis Blume, U.S. EPA, Great Lakes National Program Office
Harry B. McCarty, Judy Schofield, and Joan Cuddeback, DynCorp Systems and Solutions

The Office of Water is involved in a wide variety of activities including nationwide regulatory development, research studies, and the development of analytical methods for regulatory compliance monitoring. The Great Lakes National Program Office (GLNPO) is a geographically-focused office whose mission is to lead and coordinate United States efforts to protect and restore the Great Lakes. Both the Office of Water and GLNPO perform a wide variety of environmental data collection activities or fund those activities through other organizations. All of these data collection activities fall under the umbrella of the EPA mandatory agency-wide quality system in EPA Order 5360.1 A2.

Given the variety of data collection activities, both the Office of Water and GLNPO have adopted a graded approach to implementing quality systems. Applying a graded approach means that quality systems for different organizations vary according to the specific objectives and needs of the organization. Thus, both Offices focus on making the quality system requirements commensurate with:

- *Importance of the work*
- *Available resources*
- *Unique needs of organization, and*
- *Consequences of potential decision errors.*

The use of a graded approach can be illustrated in projects ranging from development of nationwide regulations to small grant-funded research efforts.

The mission of EPA's Office of Water is to protect the nation's water resources. To accomplish this mission, the Office of Water is involved in a wide variety of activities including nationwide regulatory development, research studies, and the development of analytical methods for regulatory compliance monitoring. The Great Lakes National Program Office (GLNPO) is a geographically-focused office whose mission is to lead and coordinate efforts by the United States to protect and restore the Great Lakes. GLNPO functions as a principal liaison with Canadian federal and provincial governments; the International Joint Commission, other Regions, the Office of International Activities, and the State Department. GLNPO works with this wide variety of stakeholders on a multitude of projects including conducting surveillance monitoring of the Great Lakes, evaluating pollution prevention initiatives, and restoring habitat and ecosystem health. Both the Office of Water and GLNPO perform a wide variety of environmental data collection activities or fund those activities through other organizations. All those data collection activities fall under the umbrella of EPA's mandatory quality system in EPA Order 5360.1 A2. Broadly speaking, the quality system requirements for environmental information collection activities performed by or funded through these Offices:

- Development of a quality system
- Documentation of that quality system
- Employing systematic planning for all environmental information collection activities, and
- Ensuring collected information is of adequate quality for the intended use.

Given the variety of these data collection activities, both the Office of Water and GLNPO have adopted a graded approach to implementing quality systems. This means that quality systems for different organizations may vary according to the specific objectives and needs of the organization. The effort needed to develop and document a quality system should be based on the scope of the program. Similarly, the effort for planning and documenting quality activities for specific projects may vary according to the nature of the work and the intended use of the data. Decisions based on environmental data vary in the level of certainty required in the underlying results. Some decisions involve a greater risk if the decision is in error, for example, the risk to public health if the level of a contaminant in drinking water is not adequately controlled. Both Offices focus on making the quality system requirements commensurate with importance of the work, available resources, unique needs of organization, and consequences of potential decision errors. The use of a graded approach can be illustrated in projects ranging from development of nationwide regulations to small grant-funded research efforts.

EPA's Coal Mining Presumptive Rule

Under the Clean Water Act, EPA establishes restrictions on the pollutants that are discharged into the nation's waterways. To do this, EPA collects and assesses information to establish regulatory limitations and standards for categories of industrial discharges. These limitations and standards are based on the degree of control that can be achieved by available, practicable, and economically feasible technologies and best management practices. Historically, establishing such standards requires that EPA engage in large primary data collection activities, complete with all the attendant quality assurance oversight of the sample collection, analysis, review, and validation procedures.

In September 1998, an effluent guidelines plan was published in the *Federal Register* (63 FR 47285) requiring that EPA re-evaluate and promulgate guidelines for the coal mining industrial category by December 2001, under an expedited and presumptive rulemaking process. The *Federal Register* Notice also described recommendations of EPA's Effluent Guidelines Task Force for expediting development of these regulations, including targeting known pollutants that are of the greatest concern and examining well-demonstrated technologies. By definition, these presumptive rules must be applicable to regulations that can be developed based on existing data, and demonstrated environmental improvements.

Toward this effort, the Office of Water initiated development of regulations designed to improve environmental conditions and pollution discharges from abandoned mine lands. The Rahall Amendment to the Clean Water Act of 1987 dictated that EPA address coal remining operations. As a result, by 1998, extensive data were available for regulatory consideration. The quality of the existing data was generally known and documented because the data were generated by States for NPDES and mining permits and compliance monitoring purposes and therefore included as part of standard discharge monitoring reports, and subjected to regulatory review.

EPA developed a series of work plans that included project schedules, roles and responsibilities, and a detailed description of data needs. EPA then coordinated with the Interstate Mining Compact Commission's (IMCC) Coal Remining Task Force consisting of EPA, the U.S. Department of Interior's Office of Surface Mining and Regulatory Enforcement (OSMRE) , and State regulatory authorities, to identify and collect data and information to be used in establishing limitations and standards. This combined effort resulted in the collection of:

- 61 data and information packages from coal mining permit applications, permit monitoring reports, and abandoned mine land remediation projects;
- 40 permit application modules;
- Data from the OSMRE and State abandoned mine land inventory databases;

- Discharge data from over 100 closed remining operations representing over 240 pre-existing discharges and more than 10 years of implementation of best management practices; and
- Responses from 19 states to an IMCC questionnaire on the status of abandoned mine lands, pollution discharges and the potential for reclamation through remining.

Throughout the evaluation and assessment of the data that were collected, the Office of Water worked closely with representatives from the OSMRE and State regulatory authorities. In order to expedite this process, the Office of Water created a database of permit and contact information, operational and production statistics, technology descriptions, site assessments, and discharge data for the target pollution parameters (i.e., iron, manganese, acidity, and solids). During completion of the database, EPA developed, conducted, and documented data entry quality control checks for more than 10% of the entire database. Information compiled into this database was used to evaluate and document the efficiency of best management practices and to establish guidelines for determination of discharge limitations. The data were used to establish requirements for determining site-specific baseline pollutant loadings in pre-existing discharges and for implementing pollution abatement plans consistent with the requirements of the Rahall Amendment (i.e., that demonstrate improvement of discharges). The discharge data from the closed remining operations (>100) were used to verify that the site-specific loadings would result in environmental improvements. In late 1998, EPA distributed a draft Coal Remining Best Management Practices Guidance Manual to regulatory authorities and other stakeholders that presented case studies and current practices of technology implementation, as well as a statistical evaluation of the efficiencies of best management practices. EPA used the statistical evaluations presented in the guidance manual to propose limitations guidelines and standards under a Coal Remining Subcategory on April 11, 2000 (65 FR 19440). Following the proposal, EPA held three public meetings to obtain comments on the proposal and the processes used to develop the regulation and responded to comments on the proposed regulations and data assessments. On January 23, 2002, EPA promulgated the Coal Remining Subcategory and published support documents describing the data collection, evaluation, and analysis activities.

The success of an effort of this scope in such a relatively short time frame would not have been possible without an effective quality system in the Office of Water that was also flexible enough to address the assessment and use of data from existing sources to produce results similar to those that have been achieved for primary data collection activities.

Great Lakes National Program Office's Lake Erie Total Phosphorus Loads

The Great Lakes contain 20% of the world's freshwater and are a globally important natural resource that are currently threatened by multiple stressors. While significant progress has been made to improve the quality of the lakes, pollutant loads from point, non-point, atmospheric, and legacy sources continue to impair ecosystem functions and limit the attainability of designated uses of these resources.

One of the best developed indicators of progress in achieving Great Lakes water quality goals is the mean annual loading of total phosphorus to each lake. For the past 25 years, particular emphasis has been placed on total phosphorus because the International Reference Group on Great Lakes Pollution from Land Use Activities recommended target loads for each of the lakes based on detailed analysis of their eutrophication status. Tributary monitoring programs and associated point source and atmospheric deposition monitoring throughout the 1980s and into the 1990s provided data to evaluate and report on progress in meeting these target loads. Unfortunately, budget cuts to environmental programs in the mid 1990s had a dramatic effect on monitoring efforts, particularly on tributary sampling. Total phosphorus load estimates were not available for any of the Great Lakes after 1995.

At the "Lake Erie in the Millennium" conference in 2001, U.S. and Canadian researchers presented the latest information on ecosystem health. Without exception, investigators reporting on water quality, plankton communities, fish stocks and the food webs that link these components, noted that their efforts have been hampered by a lack of knowledge of nutrient loads to Lake Erie. The key nutrient that drives many aspects of the Lake Erie ecosystem is phosphorus. Therefore, in 2001, GLNPO initiated a project to model phosphorus loads in Lake Erie using existing data from government agency databases including:

< EPA and State's Permit Compliance System Database

< U.S. Geological Survey's STORET Database, and

< Ontario Ministry of the Environment's Municipal and Industrial Strategy for Abatement database.

Project planning for modeling projects is important in order to ensure that the model is scientifically sound, robust and defensible. EPA's Quality Staff developed a draft checklist for use in planning modeling and other projects using existing data, *Using Data from Other Sources - A Checklist for Quality Concerns*. GLNPO's Lake Erie Total Phosphorus Loads project illustrates the use of the checklist as detailed below.

1. Identify the decision you are making or the project objectives: The overall objective of the project was to revive phosphorus load estimation efforts for the Great Lakes, using Lake Erie (1996-2000) as an example. The information required by GLNPO was the estimate of total phosphorus load for Lake Erie for the years 1996-2000. The over-arching decision was whether changes in environmental management were needed to address phosphorus loads to Lake Erie.

2. Identify the data and information from outside sources proposed for the project: A list of government agency databases was identified and used to provide data for input to study models.

3. Determine whether the data have any constraints affecting their use in the new project: The government agency databases that were used were verified to be available and accessible for use in the project.

4. Determine where the acquired data will be used in the decision making process: The data were used as inputs to the phosphorus loading model. Use of the data and model calculations were detailed.

5. Scrutinize data for quality concerns pertinent to the intended use: The Principal Investigator and GLNPO noted that the responsibility for basic data review, validation and verification was with the agency that collected the data. However, additional data screening for the purpose of load estimation was conducted prior to input to the model. The Principal Investigator used statistical programs to identify outliers by checking for internal consistency and comparing to historical information. A plan was developed to investigate outliers and determine when data would not be used in the model. Another quality concern involved the critical assumption that the quality objectives and criteria associated with the existing databases would be adequate for the purposes of this project. GLNPO and the Principal Investigator noted that numerous studies have been conducted in the past to ensure that this was the case. Data comparability also was a concern. A main project requirement was that the monitoring data be obtained in the same way as previous load estimates to ensure comparability to historical data. The Principal Investigator determined that the flow and total phosphorus data currently being generated by the agencies responsible for the existing databases are of comparable quality to data reported by these agencies previously. The Principal Investigator noted one exception and developed a plan to address this data issue.

6. Document your analysis plan in a Quality Assurance Project Plan (QAPP): GLNPO and the Principal Investigator developed a QAPP according to EPA QA/R-5, *EPA Requirements for Quality Assurance Project Plans*. In accordance with the graded approach, the QAPP noted when components discussed in R-5 were not applicable to the project. For example, Section B.4 Sampling Methods, was

cited as "not applicable" because the project used existing data and environmental sampling was not a component of the project. The plan also included a detailed description and citations for peer-reviewed equations used in the model.

7. Execute your analyses and document the outcome appropriately: The project QAPP included a schedule for model completion and development of a final report. As noted under Step 5, a main project requirement was data comparability to historical data. To address this, the Principal Investigator used standard methods for load estimation and documented any deviations in the final report.

In accordance with their quality system requirements, GLNPO implemented systematic planning for this project. EPA's draft *Guidance for Quality Assurance Project Plans for Modeling*, EPA QA/G-5M, September 2001, lists the elements that should be incorporated into the planning process. They are:

- A systematic planning process
- Peer-reviewed theory and equations
- A carefully designed life-cycle development process that minimizes errors
- Documentation of any changes from the original plan
- Clear documentation of assumptions, theory, and parameterization that is detailed enough so others can fully understand the model output
- Input data and parameters that are accurate and appropriate for the problem
- Output data that can be used to help inform decision making.

These elements were included in the planning process for the Lake Erie Total Phosphorus Loads project. In accordance with GLNPO's graded approach to quality systems, the level of effort addressing these elements was commensurate with the importance of the work and the consequences of decision errors. This project did not dictate the same level of planning and quality assurance activities that would be needed for a project directly supporting a regulation or addressing a significant threat to public health. Similarly, because the project was short-term, the level of effort addressing quality issues was not as high as a long-term project, such as GLNPO's ongoing surveillance monitoring of the Great Lakes that has a 60-page QAPP and a comprehensive sampling and analytical manual containing more than 30 SOPs. GLNPO's quality system, the Quality Staff's requirements and guidance documents, and implementation of the graded approach facilitated the success of this project to address Lake Erie ecosystem health while minimizing project expenditures.

GLNPO's Lake Erie Total Phosphorus Loads project represents the TV-dinner of quality systems - adequate and filling, but not overly involved or fancy - while the Office of Water's coal mining presumptive rule represents the filet mignon approach, complete with many courses. The coal mining presumptive rule involved a large number of cooperators and stakeholders, an extensive data search, and compilation and statistical evaluation of compiled data to generate regulatory limits. The draft checklist for planning modeling and other projects using existing data, *Using Data from Other Sources - A Checklist for Quality Concerns* also could have been used to assist in planning for this project. The graded approach to quality systems included in the Office of Water's and GLNPO's quality policies facilitates quality management supporting the wide variety of projects undertaken by these offices.

Transforming an EPA QA/R-2 Quality Management Plan into an ISO 9002 Quality Management System

Roger A. Kell, IT Corporation
Eric S. Reynolds, US EPA
Clyde M. Hedin, IT Corporation
Garabet H. Kassakhian, Ph.D., IT Corporation

The Environmental Protection Agency's (EPA) Office of Emergency and Remedial Response (OERR) requires environmental data of known quality to support Superfund hazardous waste site projects. The Quality Assurance Technical Support (QATS) Program is operated by the IT Corporation to provide EPA's Analytical Operations Center (AOC) with performance evaluation samples, reference materials, on-site laboratory auditing capabilities, data audits (including electronic media data audits), methods development, and other support services. The new QATS contract awarded in November 2000 required that the QATS Program become ISO 9000 certified. In a first for an EPA contractor, the QATS staff and management successfully transformed EPA's QA/R-2 type Quality Management Plan into a Quality Management System (QMS) that complies with the requirements of the internationally recognized ISO 9002 standard and achieved certification in the United States, Canada and throughout Europe. The presentation describes how quality system elements of ISO 9002 were implemented on an already existing quality system. The psychological and organizational challenges of the culture change in QATS' day-to-day operations will be discussed for the benefit of other ISO 9000 aspirants.

INTRODUCTION

IT Corporation operates the Quality Assurance Technical Support (QATS) Program under contract to the US Environmental Protection Agency (EPA) Office of Emergency and Remedial Response (OERR) Analytical Operations/Data Quality Center (AOC). The purpose of the QATS Program is to provide technical support to assure the collection of environmental data which are of sufficient quality to meet the technical and analytical needs of OERR and Superfund. The QATS Program provides products and services under task areas that include Performance Evaluation Samples, Tape and Data Package Reviews, On-site Laboratory Audits, Methods Evaluation, and Logistical and Administrative Support.

The QATS Program Quality System has historically operated under a Quality Management Plan prepared in accordance with EPA QA/R-2. This document is based upon the consensus standard, *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*, American National Standard Institute/American Society for Quality Control ANSI/ASQC E4-1994. In 2001, as requirements of the EPA contract awarded in November 2000, QATS completed the transition from a QA/R-2 Quality Management Plan (QMP) to a Quality Management System (QMS) that complies with the International Standard Organization ISO 9002:1994 standard.

The objective of this paper is to provide an overview of the process by which the IT QATS Program

successfully transitioned from operating under a QA/R-2 type Quality Management Plan into an ISO 9002 compliant QMS and achieved certification.

QUALITY MANAGEMENT SYSTEMS QA/R-2 has widespread application in environmental laboratories and programs. The requirements include eleven elements addressed in a Quality Management Plan (QMP). These elements include documentation of the policy, scope, applicability, and management responsibilities of the organization's quality system. The elements also include requirements for managing the system including defining the primary responsibilities for implementing each component of the system. In addition, procedures are required for ensuring adequate personnel training, procurement of items and services, controls for quality-related documents and records, and how data operations will be planned. Lastly, procedures are required for implementing work processes, determining the suitability and effectiveness of the implemented quality system, and improving the quality system.

Similarly, ISO 9000 has widespread application to many businesses, including environmental laboratories and programs. The ISO 9001 standard contains twenty elements. All of the elements of QA/R-2 are covered by the ISO 9001 standard. The additional elements in the ISO 9001 standard address the requirement for documented procedures for controlling product supplied by the customer, product identification and traceability, control of inspection, measuring, and test equipment, control of nonconforming product, etc.

ISO 9001 and 9002 are similar to each other because they contain elements that are common. The ISO 9002 standard is the same as the ISO 9001 standard but does not contain the design element. Selecting the most applicable standard depends on the type of work that is performed by the organization. The IT QATS Program chose ISO 9002 as most applicable because it contains all the elements necessary to the QATS Program. The QATS ISO 9002 QMS has advantages over the QA/R-2 Quality Management Plan, in that, (a) it requires the organization demonstrate continuous improvement of the QMS under which products and services are generated, (b) it ensures a focus on customer satisfaction, and (c) it requires less client effort to manage because oversight is provided by a third party, the ISO registrar.

APPROACH/IMPLEMENTATION The approach taken by IT to convert to an ISO 9002 QMS involved the following key components:

- Management Commitment
- ISO 9000 Consultant Selection/Gap Analysis
- Implementation Plan
- Incorporation of New/Modified Procedures
- Document Control/Tracking System
- Internal Audit Program
- Staff Training
- Management Review
- Registrar Selection

Management Commitment The primary requirement of any successful QMS implementation program is the commitment of management. When management at all levels promotes the implementation of ISO 9002, the successful accomplishment of all other subsequent tasks will follow. The key is getting staff on board. As staff continually see the commitment and enthusiasm of management to the implementation of ISO 9002, they increasingly become involved. Early on, the QATS Program Manager, the Vice President, and the President of the company communicated their commitment to the implementation of an ISO 9002 compliant quality system to all staff.

ISO 9000 Consultant Selection/Gap Analysis One of the first steps in the Quality System transition was to perform a gap analysis to compare where we currently were with a QA/R-2 compliant system and what would be required to achieve ISO 9002 certification. QATS selected a consultant with a history of

successful ISO registrations to conduct a series of three gap analyses. The recommendations provided by the consultant following each gap analysis included, but were not limited to, additional written procedures, staff training, and identification of internal auditors. An important recommendation was the promotion of the use of ISO 9002, not as a QMS only, but as a business operating system with an inherent quality. The importance of this recommendation has become clearer over time, following measurable improvements to production efficiency and quality of QATS operations.

Implementation Plan Following each gap analysis, the QATS Program developed or revised an implementation plan which included achievable milestones and time-lines. The milestones and time-lines were provided to all staff, which was key to keeping the transition on schedule. The next steps that follow were accomplished concurrently. These steps can be divided into four main areas: incorporation of new or modified procedures, document control and tracking, development of an internal audit program, and staff training.

Incorporation of New/Modified Procedures Several procedures had to be developed or modified to address the additional requirements of ISO 9002. After initiating the production of new procedures and/or modifying existing procedures, the QATS Program scheduled a series of internal audits to check the effectiveness of the QMS. Areas where shortcomings were identified were documented via Non-Conformance Reports (NCRs). Responses to NCRs required documented corrective action and subsequent follow-up and close-out by the internal auditors. Audits and subsequent NCRs are effective tools in implementing and then fine-tuning a QMS.

Document Control/Tracking System The documentation system developed at the QATS Program ensures that documents vital to the quality of the products are controlled. Only the most recent version of documents are available for use. The document system required by QA/R-2 included a Quality Assurance Program Plan (QAPP) and Standard Operating Procedures that require approval by EPA. The ISO 9002 QMS requires a four-tiered system that includes the following components:

- Level 1 - Quality Manual (New - Replaces QAPP)
- Level 2 - Quality Operational Procedures (QOPs) (New)
- Level 3 - Standard Operating Procedures (SOPs) and Work Instructions
- Level 4 - Quality Records and Other Reference Materials

The Quality Manual relates each applicable element of the ISO 9002 standard to our specific type of work, Quality Operational Procedures (QOPs) address each element of the ISO 9002 standard in detail, and Standard Operating Procedures and Work Instructions address specific procedures or activities. Quality records such as calibration records constitute the fourth level of the ISO 9002 documentation system. Other documentation, such as controlled forms, also are in the fourth level. The QATS system for control of forms and documents involves placing them electronically on the local area network (LAN). All controlled items are organized into their logical subdirectories. These subdirectories are all in one main ISO 9002 directory with limited read/write access.

After a document control system had been developed, a document tracking system was developed for procedures, forms, and all controlled documentation, including documents of external origin (i.e. documents in which the revision occurs outside the organization). Maintained by a Document Control Officer (DCO), all procedures are assigned a unique identification number and a revision number. Upon approval, new or revised procedures are provided to applicable staff and previous versions are removed. Originals of procedures (i.e. original approval signatures) are maintained by the DCO; obsolete procedures are also maintained but are marked to prevent their subsequent use. As mentioned

previously, the control system includes the NCRs which are categorized and tracked. Periodically, the NCRs are reviewed and information trended for Management Review, another important element of the standard.

Internal Audit Program One of the key elements to implementation was the development of an internal audit program. Since one of the task areas of the QATS contract is to provide support to EPA for on-site audits, the QATS Program has trained ISO 9000 Lead Auditors (nine currently) on staff. Accordingly, identification of internal auditors was a straight-forward task. However, training of internal auditors was still required to ensure they were familiar with the procedures QATS developed and implemented to comply with the standard. Similar to any audit program, the auditors observed the tasks being performed, interviewed auditees, reviewed procedures and support documentation and produced findings. The findings were reviewed and modifications to procedures were made, as appropriate. Throughout this process, auditors and staff alike gained knowledge of the QMS. As a result, performing and participating in internal audits also became a natural training tool for all staff.

Staff Training Staff training was an important part of the implementation plan. QATS staff were trained by the American Society for Quality (ASQ) in Developing and Implementing and ISO 9000 QMS and its laboratory counterpart ISO 17025. Specific staff training included ISO 9000 Internal Auditor Training an ISO 9000 Lead Auditor Training. Other training tools that were developed included flow charts for the performance of select procedures and group question and answer sessions led by management. The flow charts assisted the auditees in explaining the performance of various tasks and provided the internal auditors and the registrar with a brief explanation of the procedures being reviewed.

Management Review ISO 9002 requires periodic Management Reviews be conducted at specific intervals. Management Review involves evaluation of objective evidence from internal audits, non-conformance reports, customer satisfaction reports, and other quantitative assessments of how the QMS is measured against the objectives of quality. Each ISO 9002 element of the standard has to be internally audited prior to the first management review. A Management Review was performed prior to the registration audit with the consultant present.

Registrar Selection Selecting the appropriate registrar was also an important step to achieving ISO 9002 certification. A matrix and rating system was established as part of the selection process. In addition, the search included identification of a registrar that would be familiar with the nature of the QATS Program or operated in a similar field of expertise. After identifying the specifications for the desired qualifications and audit schedule, a request-for-quote (RFQ) was produced. Four registrars were identified as potentially having the qualifications and experience necessary. Through a documented review process, which included reference checks, a registrar was selected. The size of the facility (i.e., number of staff, number of buildings) determines number of auditors and number of days of the registration audit. For QATS, with a staff of 30, the registration audit required a three day period involving one registration auditor. The Quality Manual and QOPs were submitted to the registrar six weeks prior to the audit for review. The registration audit was conducted in October of 2001 and only three very minor corrections were necessary. Of significance during the closing meeting with the registration auditor, which was attended by all staff, the auditor noted the QATS QMS was one of the best documented systems he had seen. In addition, he was impressed with the electronic document control system. The QATS Program achieved certification to ISO 9002 on October 11, 2001.

CHALLENGES Problems and obstacles encountered along the way occurred in two categories: human factors and logistical factors. The human factors included the natural resistance to change that most people exhibit. This resistance was gradually overcome as more training was provided and knowledge of application of the standard to QATS operations increased. Another challenge was ensuring appropriate

preparation for audit questions. Most auditees initially become nervous or insecure during an audit and have difficulty providing direct responses to questions. This can blur the focus of the audit. The objective of any audit is to gather useful information and produce an audit report from which improvements can be identified. Complementing this category of challenges was a simple explanation of the terms used, such as "non-conformance" and "corrective action." People tend to take terms as "non-conformance" and "corrective action" personally so it was necessary to explain that these terms referred to the "system" and not the individual.

One of the logistical factors is the control of forms. Although form versions are controlled, some staff tend to reuse older versions. This usually occurs when an individual has electronically obtained a current form from the network and saved it as a file in their personal directory. Finding a previously completed form in one's personal directory is frequently easier than locating the blank form in a controlled directory.

BENEFITS After approximately one year of operating under an ISO 9002 QMS, the benefits support the decision to initiate implementation. As stated earlier, the benefits to ISO 9002 over QA/R-2 include requirements for the organization to demonstrate the continuous improvement of the QMS covering all products and services, a focus on customer satisfaction, and the need for less client involvement in quality system oversight since this function is now provided by the registrar. Routine surveillance audits are provided by the registrar to verify compliance to the standard. Probably the single-most important benefit is that the ISO 9002 QMS is a business operating system that possesses an inherent quality. This inherent quality is reflected in organized and easily retrievable documentation, the ability to trend processes for improvement and deficiencies, and greater attention to customer concerns. Quality is inherent because everyone is involved in the quality management system. Quality is a part of everyone's tasks and duties and is not perceived as the sole responsibility of the Quality Assurance Officer or the Quality Assurance Department. Using the system is convenient and organized; current versions of forms and procedures are easily located and trending of non-conformances allows forward-looking approaches to problem resolution.

ISO 9001:2000 On December 15, 2003, all organizations that are currently certified to the 1994 standard will be required to convert to the updated standard, ISO 9001:2000. This update has a stronger focus on customer satisfaction and less on documentation requirements than the current ISO 9000:1994 version. The QATS Program will, over the next year, convert to the new updated standard.

SUMMARY In approximately a one year period, IT Corporation effected the transition of a QA/R-2 based QMP to a QMS compliant with the ISO 9002 international standard. To our knowledge, this was the *first* such transition of this type of an EPA contract. Implementation of the ISO 9002 standard has measurably improved the efficiency of the QATS Program, and we believe will provide measurable improvements to the quality of future deliverables. An ISO 9002 QMS has distinct benefits for both suppliers and customers as compared to one based on QA/R-2. Certification to a recognized international standard requires verification by a third party, the registrar. Accordingly, an ISO 9002 QMS requires less client QA oversight to manage. In addition, the successful implementation of an ISO 9002 QMS emphasizes continuous improvement of the quality system while focusing on customer satisfaction. Although these elements were present and active prior to implementing an ISO 9002 QMS, after implementation, they are supported by all staff and all activities and, therefore, comprise a business operating system that possesses an inherent quality.

REFERENCES

1. ANSI/ISO/ASQC Q9001-1994, *Quality Systems-Model for Quality Assurance in Design, Development, Production, Installation, and Servicing,* American Society for Quality, Milwaukee, Wisconsin 53210-3005 (1994).

2. ANSI/ISO/ASQC Q9002-1994, *Quality Systems-Model for Quality Assurance in Production, Installation, and Servicing,* American Society for Quality, Milwaukee, Wisconsin 53210-3005 (1994).

3. ISO (International Organization for Standardization), *International Standard ISO 9001:2000 (E) Quality Management Systems-Requirements,* Third Edition, ISO, Geneva, Switzerland (2000-12-15).

4. United States Environmental Protection Agency, *EPA Requirements for Quality Management Plans, EPA QA/R-2,* EPA/240/B-01/002, Office of Environmental Information, Washington, DC 20460 (March 2001).

5. Reference Website: http//www.iso.ch

Enforcement and Compliance Data Quality Strategy

David Sprague, U.S. EPA

The Enforcement and Compliance Data Quality Strategy's (DQS) purpose is to develop and implement an ambitious, practical plan to assess and manage the quality of enforcement and compliance data. The vision of the DQS is to assure that the data used to support enforcement program decisions are of high quality and accurately reflect the activity and accomplishments of federal and state compliance and enforcement programs. High quality data is defined as accurate, complete, and timely data that are clearly presented and consistent with national data standards. The Agency can confidently disseminate and use this information to evaluate programs, target resources and to inform Congress, the public, and the news media.

The Strategy is designed to reshape ad hoc data quality efforts and to establish routine and systematic methods for improving and maintaining a high level of data quality. The focus of this strategy will be to identify problems by examining enforcement and compliance data in Agency's systems and raising issues to the appropriate personnel for further analysis and correction. In addition, ongoing activities, such as the development of online correction, will be undertaken at the same time to address data quality problems that are long-term projects or may require activities on repeating schedules (e.g., training).

I. COMMON DATA QUALITY ISSUES

The data quality activities discussed in this Strategy are specifically aimed at identifying data that is missing, incorrect, or inadequate. Definitions of these terms are provided below.

Missing Data refer to required files, records or values that are not in the EPA national databases due to incomplete reporting by EPA Headquarters, EPA Region, a state agency, a local authority, or facility. This deficiency precludes accurate program evaluations and could suggest that important environmental work is not being done.

Incorrect Data are values within a field that do not accurately represent the true value. There are three main reasons for data in a required field to be incorrect. These reasons are data creation errors, data entry errors, and ambiguity in use of data fields.

Inadequate Data are defined as records or fields that are not accessible or usable as currently maintained in EPA national databases. For example, data fields might exist in systems, but EPA does not require reporting or has allowed the fields to fall into disuse. Alternatively, new data needs may have arisen due to programmatic changes and the data systems were not modified to meet those needs. In other instances, inadequate data may result if regulations or policies have not been updated to match new goals, technologies, or systems.

II. DATA QUALITY ACTIVITIES

This Data Quality Strategy proposes various activities to identify and correct current data quality problems. Annually, the Office of Compliance (OC) will develop an enforcement and compliance Data Quality Strategy Implementation Plan in consultation with Regions and states. The plan's goal is to identify specific data quality issues through periodic activities, such as audits and analyses. The outcome of these activities will be used to prioritize the order in which to address any data problems. Resulting analyses will be sent to Regions and states to review and correct. Ongoing activities, such as the

development of online correction, will be undertaken at the same time to address data quality problems that are long-term projects or may require activities on repeating schedules (e.g., training).

Proposed activities are organized and described in this section according to whether they are periodic or on-going. Figure 1 shows how the activities fit within the DQS and how they are expected to complement each other.

A. Periodic Activities

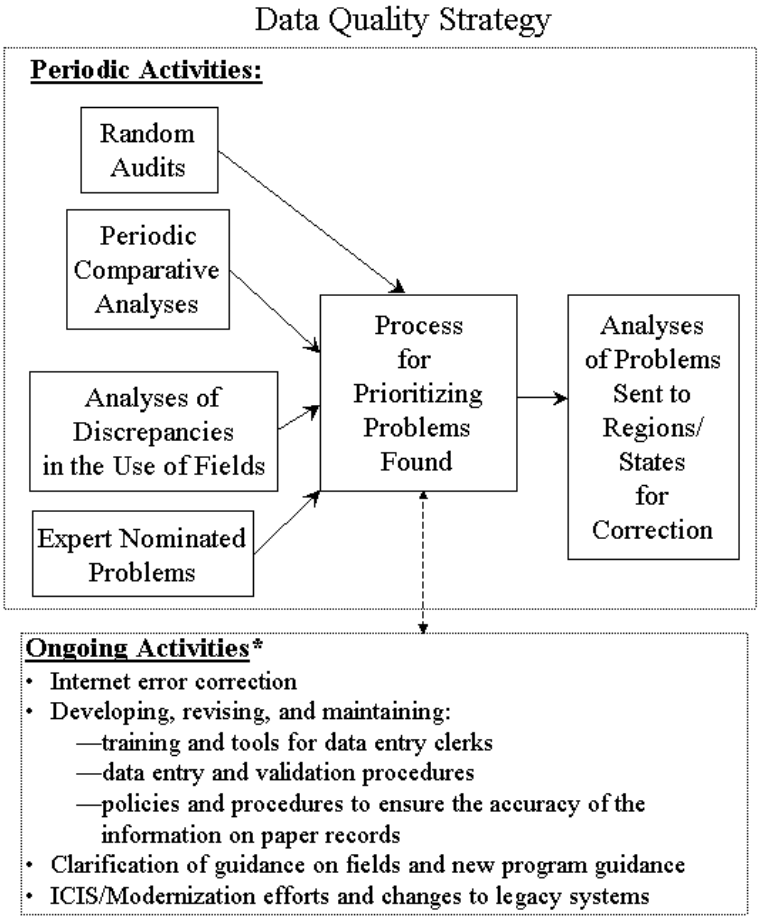
OC will develop an annual Implementation Plan in consultation with the Regions and states. The will identify specific projects for the coming year. It will include details and planning information for identifying data errors, prioritizing problems, selecting data quality projects and involving Regions and states; and sending problem analyses to Regions and states for correction.

Identifying Data Errors

Data problems will be identified through methods, such as: random data audits; comparative analyses; analyses of discrepancies in the use of fields; and expert nominated problems.

- Random Data Audits. Data in selected key data fields will be independently evaluated through random data audits to validate its accuracy and completeness.
- Periodic Comparative Analyses. Periodically, comparative analyses of particular data fields across organizational units with delegated authority will be conducted to identify potential data quality problems. For example, comparing Significant Non-Compliance or inspection rates within Regions or states to the national average can be used to identify extreme outlying values that might indicate data problems.
- Analyses of Discrepancies in the Use of Fields. OC will analyze key enforcement and compliance activity data fields to determine if there are discrepancies in Regional or state usage. Discrepancies found will be documented, and guidance developed to assist program implementers and database users in how to use the codes for nationally consistent reporting.
- Expert Nominated Problems. Staff with significant experience/expertise with individual data systems (data system staff, program staff or expert users) will identify all quality issues they encounter. It is expected that such data quality problems will be identified from intensive usage, such as targeting and measures analyses. Data problems found by comparing data required to be entered into more than one database can be used to reconcile information across data systems.

Figure 1: Activities Established by the Data Quality Strategy



*Any of these ongoing activities may uncover problems that need to be addressed formally. Solutions for some prioritized problems may result in revisions or new data entry procedures, clarification of guidance, or system changes/modernization.

Prioritizing Data Problems Found

Data problems will be prioritized by a standing data quality workgroup that will consult with relevant stakeholders. Data quality efforts will focus on:

< data in the legacy database that will next be modernized;

< administrative priorities of stakeholder agencies and offices involved in management of each data system;

< EPA administration priorities;

< areas of concern raised by the Environmental Council of the States (ECOS), the Inspector General, EPA Management, Regional information users and managers;

< areas of concern discovered by methodologies for identifying data problems;

< areas of concern raised by public access; and

< Regional and state review of assessments that confirm validity of the problems.

Selecting Data Quality Projects and Involving the Regions and States

Development of the annual DQS Implementation Plan will follow a structured decision making process. This process will resemble the one used for development of the FY02 Implementation Plan and is described below. Each year, the Office of Enforcement and Compliance Assurance's (OECA) MOA Guidance will specify national data quality projects for EPA and states to work on during each year.

The process will start with an annual Data Quality Strategy workshop held in the Spring. Participants will include staff from headquarters, Regions and state representatives who will:

< discuss issues related to the identification and implementation of the current year's projects; and

< brainstorm ideas for the next year's projects.

Sub-workgroups will be convened to build on the discussions from the DQS workshop and to flesh out the projects to be recommended for implementation. During the summer, the sub-workgroups will make their recommendations, which will then be sent to the entire DQS workgroup for review and comment.

The detailed implementation plan for the upcoming year will then be drafted and distributed to a wider set of stakeholders. This implementation strategy will be distributed for review and comment to the entire data steward community, before being finalized. Information on planning, assignment of responsibilities, progress, and results will be coordinated with Regions and states, via the Regions. State and Regions will have scheduled opportunities to communicate their comments on the projects to Headquarters.

Sending Analyses of Problems to Regions and States for Correction

When the data quality problems are identified via Periodic Comparative Analyses, Analyses of Discrepancies in the Use of Fields, and Expert Nominated Problems and assessments are complete, and problems are researched and well documented, Regions and states will be informed of the problems through distribution of analysis reports in various formats (e.g., short DQ alerts, longer memoranda). These reports will describe in detail an identified data quality problem narratively, and, where applicable, graphically and quantitatively. They will be written for the selected problems and sent to the responsible Regions, states (via the Regions), and Headquarters personnel to alert them of potential data problem so they can correct the data already in the system and make any other changes need to avoid the problems in the future.

B. Ongoing Activities

The following activities will be undertaken on an ongoing basis to uncover, prioritize and address data quality problems: Internet error correction; promoting accuracy of initial reporting, paper records, data entry, and validation procedures; updating guidance documents for activities that affect data quality; and updating data system documentation, support system modernization development.

Many activities will be influenced by and will influence the periodic activities described in the previous section (Figure 1). For example, data errors identified through periodic audits and analyses may spur changes to ongoing data quality work (e.g., revisions to data entry procedures, clarification of guidance, or system changes/modernization). In turn, the ongoing data quality activities will directly impact the types and number of data quality concerns.

Internet Error Correction

Currently available Internet error correction tools will be applied to particular records in need of correction. For example, OC will use information from the Headquarters, Regional, and state data steward networks, as well as the On-line Targeting and Information System (OTIS) site (and the Public Access Internet site when it becomes available) and its error correction process to pin-point data records within OC's national databases in need of correction. The Office of Environmental Information (OEI) is expected to improve the on-line system for informing data stewards at the Regional and state level about potential problems with individual records reported by data system users. OEI's system will also track

the performance of Regional and state data stewards responsible for making corrections. Resources and encouragement from program management will be critical to ensure the continued and effective involvement of those data stewards responsible for responding to error correction requests.

Promoting Accuracy of Reports, Paper Records, and Data Entry and Validation Procedures

There needs to be a commitment from Regions and states to enter required data into EPA's national databases in a timely and accurate manner. The following are activities OC may undertake, or encourage the Regions and states to undertake, to ensure that data quality procedures are accurately documented and followed.

< Utilize and document existing edit and validation checks in the legacy databases, and make recommendations for new edit and validation checks to be incorporated as the databases are modernized. As part of OC's Quality Management Plan, Data System Quality Assurance Plans will be developed for each data systems where OC has primary responsibility. The Data System Quality Assurance Plans will describe the existing and planned editing checks and validation procedures that are part of each data systems.

< Develop new policies and procedures to minimize data errors stemming from the filling in of field reports and paper records (e.g., inspection reports, enforcement action reports, case conclusion data sheets).

< Provide training for all staff generating and entering data that ultimately resides in EPA's data systems.

< Inform Regional and national program management of the need for sufficient resources to support the implementation of more comprehensive and accurate data entry and verification procedures.

< Conduct a one-time review of instructions for data submission to ensure that Regions and states provide data in formats consistent with data system requirements (e.g., values and/or formats should be consistent with data system or translatable with full documentation).

Updating Data System Guidance

Data system guidance and documentation will be updated to address discrepancies in the use of key enforcement and compliance activity database fields. Fields inconsistently used by Regions and states could stem from different understandings of how database fields are to be used, or from outdated media-specific program guidance.

System Updates and Modernization

OC will work to maintain current, well documented system data dictionaries. Such efforts to create and maintain all required documentation will be explicitly called for as part of OC's Quality Management Plan. These data dictionaries should provide clear definition of data fields and, when necessary, clear definitions of programmatic terms, such as, what a "final judicial order" means for the RCRA program.

OC will also work with the modernization team to ensure that the data quality improvements recommended by the Data Quality Strategy are built into the modernized data systems.

III. REGIONAL AND STATE IMPLEMENTATION OF THE STRATEGY

Since Regions and states are responsible for entry of enforcement and compliance information, their active participation is critical. OECA's Memorandum of Agreement Guidance will continue to specify the Region's level of commitment to implementing the DQS.

OC will propose the following efforts to ensure Regions and states implement the activities of this Data Quality Strategy:

< Provide yearly quality assurance awards to states.

- Do not give credit for program actions and results that are not entered in a national database. Notify programs, Regions and state staff and management of this policy.
- Make any OECA discretionary extramural funding given to states dependent on whether the state commits to and prioritizes maintaining quality data. Maintaining quality data includes providing the data in a form consistent with EPA data system requirements and codes.
- Highlight data quality issues in the Regional profiles that are provided to senior management.
- Publish data quality statistics for Regions to promote positive peer pressure.

Relationship to Quality Management Plans

Although Region's have their own Quality Management Plans (QMP) to ensure data quality, this Enforcement and Compliance Data Quality Strategy covers areas not traditionally covered by Regional QMPs. This fiscal year, OC is also updating its Office-level Quality Management Plan (QMP), which is intended to clearly and fully document the policies, work processes, resources, management structure, and other critical elements of OC's data quality program. The updated QMP will detail how components of the DQS, such as data quality assessments, will inform ongoing planning and implementation of data quality activities and data system management.

EPA Information Quality Guidelines

Evangeline Tsibris Cummings, Office of Information Analysis and Access

In December 2001, Office of Environmental Information (OEI) Assistant Administrator and EPA Chief Information Officer, Kim Nelson invited agency wide participation in developing the EPA Information Quality guidelines. Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law 106-554) directed OMB to issue government-wide guidelines that "provide policy and procedural guidance to Federal agencies for ensuring and maximizing the quality, objectivity, utility, and integrity of information (including statistical information) disseminated by Federal agencies." The OMB guidelines were first issued on September 29, 2001, and are summarized below. They can be viewed in their entirety at http://www.whitehouse.gov/omb/fedreg/reproducible.html.

Statutory requirement:

Section 515 of the Treasury and General Government Appropriations Action for FY01 (Public Law 106-554) Directs OMB to issue guidelines that provide policy and procedural guidance to federal agencies and that require federal agencies to:

- Issue guidelines for ensuring and maximizing the quality objectivity, utility and integrity of information disseminated by the agency
- Establish administrative mechanisms allowing affected persons to obtain correction of information that does not comply with the guidelines
- Report periodically to OMB on the number and disposition of complaints
-

The OMB guidelines (issued first on September 29, 2001) direct agencies to:

- Adopt a *basic standard of quality* as a performance goal and incorporate the standard into Agency operations;
- Develop a process for reviewing the quality of information before it is disseminated;
- Establish mechanisms for “affected persons” to request correction of disseminated information that does not comply with the guidelines; and
- Report to OMB the number and nature of complaints we receive regarding Agency compliance with the guidelines and how complaints were resolved.

OEI will lead development of *EPA Information Quality Guidelines* and will report the Agency's guidelines as shown in the table below.

OMB timeline	***EPA TASKS***
May 1, 2002	Publish notice of availability of the draft EPA Information Quality guidelines in the *Federal Register* to facilitate public comment. Post EPA Information Quality guidelines on the EPA web site.
May 1 - June 1	Public comment period. EPA public meeting to be held in mid-May in Washington, DC.
July 1, 2002	Revised report is due to OMB for OMB Review.
October 1, 2002	Publish notice of availability of EPA's FINAL Information Quality guidelines in the Federal Register. Post EPA FINAL Information Quality guidelines on the EPA website. *(October 1, 2002 is the statutory deadline)*

Contacts: Evangeline Tsibris Cummings (202-260-1655) of the Office of Information Analysis and Access and Jeffrey Worthington (202-564-5174) of the Office of Planning, Resources and Outreach are serving as OEI's co-chairs for this work effort. *The most up-to-date information will be presented by Evangeline Tsibris Cummings at the April '02 Quality conference*.

Understanding Enterprise Information Architectures for Managing Quality

Mark Doehnert, Quality Assurance Manager
U.S. EPA Office of Radiation and Indoor Air

The enterprise information architecture is crucial to ensure the quality of information produced, managed, distributed, and exchanged by an enterprise. The author briefly reviews the basics of enterprise information architectures, applying architectural principles to information systems for an enterprise and the relationships to business planning and quality systems.

Organizations today, both public and private, rely more and more heavily on information management and information technology (IT), especially in our service-oriented economy. A recent issue of Business week Online points out that information technology now represents more than 50% of all business equipment spending vs. less than 20% some 25 years ago, and that information technology "is so crucial to business operations today -- and so expensive -- that Chief Executive Officers (CEOs) have no choice but to understand it."[1] According to CIO (Chief Information Officer) Magazine, information technology (IT) "is now a catalyst that triggers massive, beneficial transformations in organizations, markets, industries and even the world. It changes the way companies do business." IT has progressed from supporting a business, to enabling a business, to really being a driver of business.[2]

Quality is as important in information management and IT with service industries as it is in manufacturing. Quality expert W. Edwards Deming points out that one finds in most service industries large volumes of transactions and paper, large amounts of processing, and an extremely large number of ways to make errors.[3] Today the large volumes of paper also include large volumes of electronic data. The American Society for Quality (ASQ) states that quality is aimed at performance excellence and increased customer satisfaction. Quality also reduces cycle time and costs, and eliminates errors and rework. According to ASQ, quality is an approach to business using a collection of powerful tools and concepts that are proven to work.[4]

A quality system is the means by which an organization manages its quality aspects in a systematic, organized manner and provides a framework for planning, implementing, and assessing work performed by an organization and for carrying out required quality assurance and quality control activities.[5] Quality managers should understand how to manage quality in the

---

[1]Eric Wahlgren. "The Digital Age Storms the Corner Office." Business week online. September 6, 2001. www.businessweek.com/technology/content/sep2001/tc2001096_253.htm

[2] Katherine Noyes. "Enterprise Value Awards." CIO Magazine. February 1, 2002.

[3]W. Edwards Deming. Out of the Crisis, MIT Center for Advanced Engineering Study, Boston, MA.1982.

[4]"Why Quality?" American Society for Quality Internet www.asq.org/info/

[5]U. S. EPA Quality Internet site http://www.epa.gov/quality/faq2.html

information management and information technology arena just as they manage quality in more traditional areas such as measurement and manufacturing. This includes understanding how to apply the quality system to quality management in areas such as enterprise architectures, data standards development, software and software engineering quality, hardware quality, geospatial data and applications, information security, and applicable standards. As a start, this should include understanding concepts and tools like enterprise architectures and enterprise architecture planning.

One definition of Enterprise Architecture (EA) is the "explicit description and documentation of the current and desired relationships among business and management processes and information technology. It describes the "current architecture" and "target architecture" to include the rules and standards and systems life cycle information."[6] John Zachman, an expert on enterprise architecture, defines an architecture as "that set of design artifacts, or descriptive representations, that are relevant for describing an object such that it can be produced to requirements (quality) as well as maintained over the period of its useful life (change)."[7] An architecture is also defined as "the fundamental organization of a system embodied in its components, their relationships to each other and to the environment and the principles guiding its design and evolution[8]." The EA defines principles and goals and set direction on such issues as the promotion of interoperability, open systems, public access, end customer or knowledge worker satisfaction, and information security.[6]

Like quality systems, to be successful, enterprise architecture efforts are best driven from the perspective of business managers, not by the information technology managers. Author Melissa Cook says "business leadership for EA development is a must because only the business leaders understand the true information processing needs of the enterprise. It is also a must because it will require executive level understanding and commitment to manage the conflicts that inevitably occur when moving . . . to a controlled and coordinated approach."[9] Another EA EAP expert, Dr. Steven Spewak, points out " the mission of information systems is to provide quality data to those who need it. Quality and business managers can gain much from understanding what an EA is, especially because we need to manage quality for information projects just as we do for environmental data collection projects."[10]

---

[6]U. S. Office of Management and Budget Circular No. A-130, Management of Federal Information Resources www.whitehouse.gov/omb/circulars/a130/a130trans4.html

[7]John A. Zachman. "Enterprise Architecture: The Issue of the Century." May 1997 issue of Database Programming and Design. 1996.

[8]IEEE Std 1471-2000, Recommended Practice for Architectural Description of Software-Intensive Systems.

[9]Melissa A. Cook. Building Enterprise Information Architecture: Reengineering Information Systems. Prentice Hall PTR, Upper Saddle River, New Jersey. 1996.

[10] Steven H. Spewak and Steven C. Hill. Enterprise Architecture Planning - Developing a Blueprint for Data, Applications and Technology. John Wiley and Sons, New York. 1992.

Enterprise architectures are needed for complex engineering products and systems and for complex product development and construction processes. They are needed for products and systems built piece by piece, part by part, system by system, program by program and for products. Examples of such complex products and systems include the systems of databases, storage devices, computers like desktop and server, software, network routers, and interfaces present today in organizations. As systems age, changes such as upgrades are needed, so systems that have a long life time need architectures. Enterprise architectures are needed for the same reasons that we architect products like houses and airplanes, to ensure quality, to move from strategy to implementation, to ensure integration so that different parts to all fit together, to manage change from the baseline, to manage change, to reduced time to delivery or market, to reduce costs, and to provide documentation.

Architectures are also needed because of the stovepipe processes and systems that lead to unnecessary duplication and redundancy. Organizations are faced with heterogeneous, incompatible systems that limit interoperability and information exchange. Changes to data and information and IT systems are difficult to manage. Without effective architectures, an organization cannot share information, cannot collaborate, and cannot communicate. These architecture-related problems lead to increased cost, rework, waste, bad decisions, and even conflict. The need for information assurance, such as security, also drives the need to effective information architectures.

Just as the design of a measurement project or manufactured item is critical to the quality, the design of the information architecture and the pieces within that architecture are critical to information quality. There are now tools and concepts that can be applied to information architectures. Larry English, an expert on information quality, introduces the concepts of "data definition and information architecture quality," "data content quality," and "data presentation quality." For example, information architecture quality represents how well the data structure in an architecture represents real world objects and events.[11] Information quality experts also provide tools, such as Michael Brackett's evaluation criteria for achieving data resource quality.[12]

For the quality manager, ensuring that the quality system is the means by which an organization manages its information quality aspects in a systematic, organized manner requires that the quality manager understand and become involved with the organization's enterprise architecture efforts. The same kind of planning, implementing, and assessing work performed for traditional measurement activities are applied to information management activities, including software development. Finally, quality assurance and quality control activities must be planned and carried out on information architecture and software projects, all in a graded approach.

The quality manager should also be involved with other organization business process improvement efforts that directly relate to the quality system and the information architecture, such as applying the Carnegie Mellon Software Engineering Institute (SEI) Capability Maturity

---

[11]Larry P. English. Improving Data Warehouse and Business Information Quality. John Wiley and Sons. New York. 1999.

[12]Michael H. Brackett. Data Resource Quality, Turning Bad Habits into Good Practices. Addison Wesley. Boston, MA. 2000.

Models (CMMs), which are integrated process improvement reference models to advance the state of the practice of software engineering and to improve the quality of systems that depend on software.[13]

[13]Carnegie Mellon Software Engineering Institute (SEI) Internet site www.sei.cmu.edu/about/about.html

Atlanta Supersite Quality Assurance Final Assessment Report

Dennis K. Mikel, EPA - Office of Air Quality, Planning and Standards
Emission, Monitoring and Analysis Division, Monitoring and Quality Assurance Group
MD C339-02
Research Triangle Park, North Carolina, 27711
919-541-5511
mikel.dennisk@epa.gov

This paper outlines the assessment of data collected during an air monitoring research intensive, dubbed the "Atlanta Supersite," and which was conducted August 3-31, 1999. The Atlanta Supersite was one of the first research projects for speciated particle matter (PM), ozone and ozone percursors in the country. The author was the Quality Assurance Manager for the project while working for EPA Region 4 in Atlanta, Georgia. This report outlines the methods which the author used to assess the quality of the data and illustrates the results of the analysis.

1. Description

The "Atlanta Supersite Field Experiment" was conducted between the dates of August 3 – 31, 1999 in Atlanta, Georgia. This research project was conceived and implemented by a number of university, private contractor and U.S. Environmental Protection Agency researchers. The Atlanta Supersite was the first of its kind; the first time that fine particle research instruments had been brought together into one location with well established fine particle Federal Reference Methods, and instruments that analyze for photochemical precursors and oxidants. There were a several instruments that were operated for the first time in a field setting.

The following statements can be made about the quality of the data set:

- The accuracy audit data shows that the audited instruments were accurate when compared to audit standards that were administered by the EPA - Region 4 laboratory.

- The data completeness (study average was 87%) goal of 75% was exceeded. However, the data completeness for the meteorological parameters was 72.6%, which was less than the data completeness goal.

- Only a small portion of the researchers submitted precision data for the study. The author believes that the bias and comparability data are better estimates of the uncertainty of the data. The comparability data estimates whether the data are normally distributed about the mean with a confidence of 95%. In most cases, the data are normally distributed.

- A major portion of the researchers did submit minimum detection limits data.

- The bias data illustrates that the majority of the elemental parameters are within the target goal of +/- 25% with the exception of several samplers.

- The bias results for the Organic Carbon (OC), Elemental Carbon (EC) and Nitrates illustrate that the majority of samplers are outside of the +/-25% target goal. The bias data shows that the ammonium and sulfates analyses are within the target goal.

- The comparability and bias data show a very strong negative bias for the filter based EC data. This trend is the opposite with the OC data. The filter-based systems OC data show a strong positive bias.

- The bias results for the gaseous formaldehyde, Nitrous ion (HONO) and oxalate illustrate that these data are outside of the target goal of +/- 25%.

- Ozone, sulfur dioxide, nitric oxide, reactive NOy bias data shows that these parameters are within the bias target of +/- 25%.

- The monitoring location was in an excellent location in term of representativeness and exposure. The monitoring site was located in an industrial sector of the city of Atlanta, Georgia. Representative scale was determined to be urban for ozone and neighborhood scale for fine particles.

2. Project Description and Data Analysis Overview

The "Supersite" program was first conceived as a set of special studies extending beyond national regulatory network for fine particles to elucidate source-receptor relationships and atmospheric processes in support of State Implementation Plans. The program would establish monitoring centers in 4-7 airsheds representing a spectrum of PM problems across the country. Spurred by the recommendations of the National Academy of Sciences committee on PM research, EPA staff further developed the mission of the Supersite program to address priority health and exposure related research needs identified by the committee through a coordinated monitoring/coordinated science planning effort. An important part of the effort was instituting a dialogue among health and atmospheric science disciplines and research and regulatory groups, such as took place at the July, 1998 workshop on PM Measurements held in Chapel Hill, North Carolina.

In recognition of the growing concern over the deleterious health effects of atmospheric fine particulate matter and the commonalties and synergism that exist between photochemical oxidants and Particulate Matter 2.5 micron ($PM_{2.5}$), the Southern Oxidant Study (SOS) began making a transition in late 1997 from a research and assessment program concerned primarily with ozone and other oxidants in rural and urban areas of the South, to a research and assessment program concerned with fine particles. Shortly thereafter, SOS began planning for a major field experiment during the summer of 1999 to address key scientific issues related to the interactions and couplings between the formation of photochemical oxidants and . EPA decided that Atlanta would be the center for one of two initial Supersite Programs (the other one being located at Fresno-Bakersfield, California). In December 1998, the SOS Science Team was contacted by officials from the EPA and requested that it develop a plan for the Atlanta Supersite that could be implemented during the Fiscal Year 1999-2000.

In August 1999 many emerging and/or state-of-the-science measurement methods for fine, airborne particles were deployed at a site in Atlanta, Georgia, from the period of August 3, through 31, 1999. These measurements were made as part of the first of the regional Supersites being established. The Atlanta Supersite was coordinated by the SOS in collaboration with the numerous universities and agencies that comprise SOS as well as a number of other programs and agencies including the Southeastern Aerosol Research Characterization/ Aerosol Research Inhalation Epidemiology Study (SEARCH/ARIES) and SCISSAP.

3. Data Quality Objectives

The Measurement Quality Objective (MQO) indicators for the Atlanta Supersite Experiment were defined in the Quality Assurance Project Plan. The MQO indicators used in this report are listed below.

- Accuracy;
- Precision;
- Minimum Detection Limits (MDLs);
- Bias;
- Comparability;
- Completeness;
- Representativeness.

After the field experiment was finished, the Quality Assurance Manager set out to assess the data and setting the results in a Quality Assurance Final Report (QAFR). The QAFR attempts to quantify the error of the data generated. This was accomplished by utilizing the performance audits on gas phase instruments, accuracy flow checks on filter based and semi-continuous particle instruments, Technical System Audits (TSAs) and statistical tests. The QA data collected by the QA Team were used to document accuracy. Data generated by the researchers were used to determine the MDLs and precision (where available and submitted). The bias, comparability and completeness data are generated using standard statistical tests.

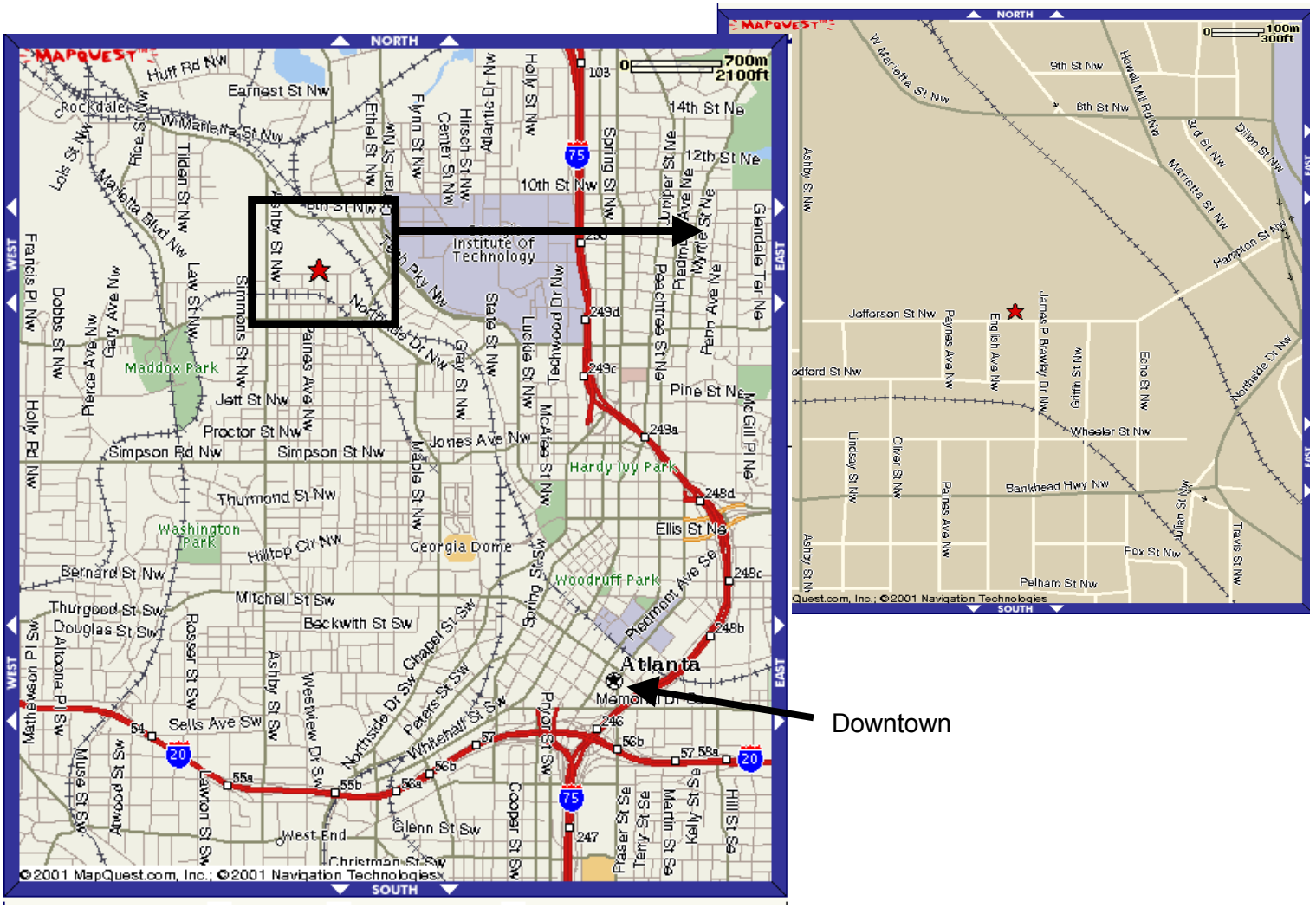


Figure 1. Map of the Atlanta Supersite Field Experiment

Figure 2. Overhead View of the Atlanta Supersite Field Experiment

Speciation Trends Network

Dennis K. Mikel, Air Quality, Planning and Standards, Emission,
Monitoring and Analysis Division, U.S. EPA

The paper will outline the Quality Assurance (QA) Program that has been implemented for the fine particle matter (PM) Speciation Trends Network (STN). The STN supports the Federal Reference Method fine particle 2.5 micron ($PM_{2.5}$) program that has been implemented since 1998. This paper will show that the QA program is not initiated by only one agency; this is a program that is managed by OAQPS but is being instituted by several programs with EPA.

1. Introduction

The Clean Air Act (CAA) requires EPA to revise or update the air quality standards based on review of the latest scientific information on known and potential human health effects associated with Particulate Matter (PM) levels found in the ambient air. In fulfilling the obligation of the law, the EPA recently reviewed the air quality criteria, National Ambient Air Quality Standards (NAAQS) for PM and epidemiological evidence that shows an association between ambient concentrations of PM and a range of serious health effects. Based on the results of its review, the EPA revised and promulgated two new primary standards for the fine fraction of PM and the regulatory requirements for monitoring the chemical composition of these particles. In response to this promulgation, EPA has instituted a $PM_{2.5}$ network. Please see Figure 1, which illustrates the overall national fine particle network. As can be seen from the this figure, the second tier of the pyramid deals with the routine speciation.

In meeting the requirements to monitor and gather data on the chemical makeup of fine particles, EPA is establishing a Speciation Trends Network. These STN samplers will be placed at various national air monitoring stations (NAMS) and State and local air monitoring stations (SLAMS) across the Nation. It is currently anticipated that 54 of these chemical speciation sites will be used to determine, over a period of several years, trends in concentration levels of selected ions, metals, carbon species, and organic compounds in $PM_{2.5}$. Further breakdown on the location or placement of the trends sites requires that approximately 20 of the monitoring sites be placed at existing Photochemical Assessment Monitoring Stations (PAMS). The placement of the remaining trends sites will be coordinated by EPA, the Regional Offices, and the state and local agencies.

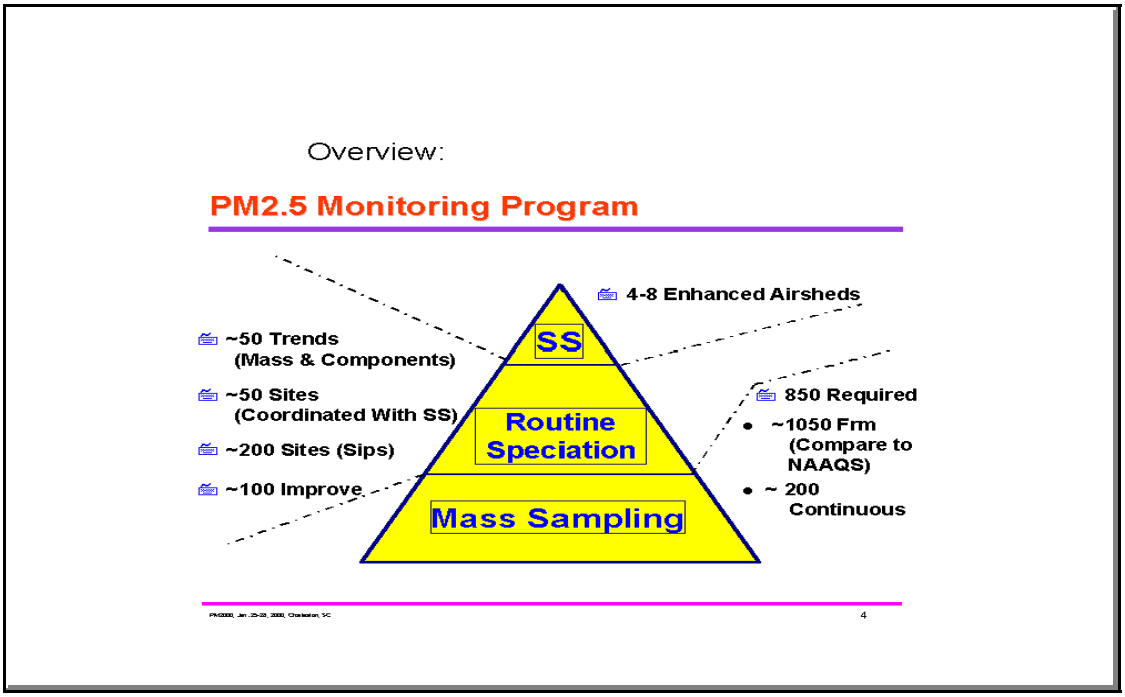


Figure 1. Overview of the National Fine Particle Network.

Locations will be primarily in or near larger Metropolitan Statistical Areas (MSAs). The remaining chemical speciation sites will be used to enhance the required trends network and to provide information for developing effective State Implementation Plans (SIPs).

As Figure 1 illustrates, the STN is a component of the National $PM_{2.5}$ Monitoring Network. Although the STN is intended to complement the activities of the much larger gravimetric $PM_{2.5}$ measurements network component (whose goal is to establish if NAAQS standards are being attained), STN data will not be used for attainment or non-attainment decisions. The programmatic objectives of the STN network are:

- *Annual and seasonal spatial characterization of aerosols;*
- *Air quality trends analysis and tracking the progress of control programs;*
- *Integration of chemical speciation data set with the data collected from the IMPROVE network; and*
- *Development of emission control strategies.*

Stakeholders in the STN will be those at EPA and State and Local agency investigators who are seeking to determine concentration trends of $PM_{2.5}$ chemical species over a period of 3 or more years and decision-makers at state and local levels who will use the data as input to models and for development of emission control strategies and determination of their long-term effectiveness. Other users will be public health officials and epidemiological researchers. However, expectations for data sets from the STN must be put in context. A number of limitations are recognized, (for instance, the 24-hour integrated sample approach, taken every 3rd day, is not suitable for determination of diurnal patterns and may have limited use to those who study health effects). EPA recognizes these data use limitations and limitations on the sampling and analysis methodologies. Thus, EPA does not rule out the possibility that objectives, requirements, and methods for speciation sampling may need to be adjusted in the future.

EPA anticipates that approximately 250 sites will comprise the full chemical speciation network. In addition to the 54 sites for the trends network, another 200 sites will be implemented to enhance the required network and provide information for developing effective SIPs. The non-trend sites will be allowed flexibility in terms of sampling frequency, site selection, site mobility, and addition of target species to address regional and local issues as needed. For example, some areas may choose to focus on specific episodes or seasons, such as a winter time wood smoke problem. EPA does not believe that a single nationwide approach to speciation sampling and analysis is the best approach for all locations. The EPA expects that most sites will follow a sampling and analysis program similar to the core STNs for their non-trend sites; however, alternative approaches will be considered on a case-by-case basis through negotiation by State agencies with EPA Regional Offices and Headquarters. EPA encourages State and Local Agencies to consider additional chemical analyses beyond the constituents specified for STN. For example, detailed analysis for compounds comprising the organic carbon fraction could provide valuable insight into development of more refined source-receptor relations, particularly in areas with significant carbon based aerosols. EPA also encourages the use of continuous monitoring techniques to the extent possible. Recent advances in measurement technologies have been proven reliable.

2. Roles and Responsibilities

OAQPS-EMAD: At the top of the QA structure is the OAQPS-Quality Assurance Coordinator (QAC). It is the QAC's responsibility to oversee that QA is implemented into the program. The QAC will interact with a QA Workgroup that has formed. This QA workgroup consists of EPA, State and local agency, Regional Office, R&IEL and NAREL QA staff. They will meet periodically to discuss QA issues as they arise throughout the program. The QAC will also work directly with the ORIA offices. Assessment Reports will be given to the QAC on an annual basis. These will include the results of the assessments listed in Table 2.1 performed during the previous year. The ORIA offices; NAREL and R&IEL will have important roles in the QA system.. In addition, the MQAG may also perform MSR or TSAs on any of the agencies in the QA or monitoring system.

NAREL-Montgomery, Alabama: NAREL will have QA oversight of the RTI laboratory operations. As such, the laboratory will perform TSAs or MSRs on the RTI laboratory operation on an annual basis. In addition, the NAREL will create laboratory audit samples Performance Evaluations (PE) that will be forward to the RTI lab.

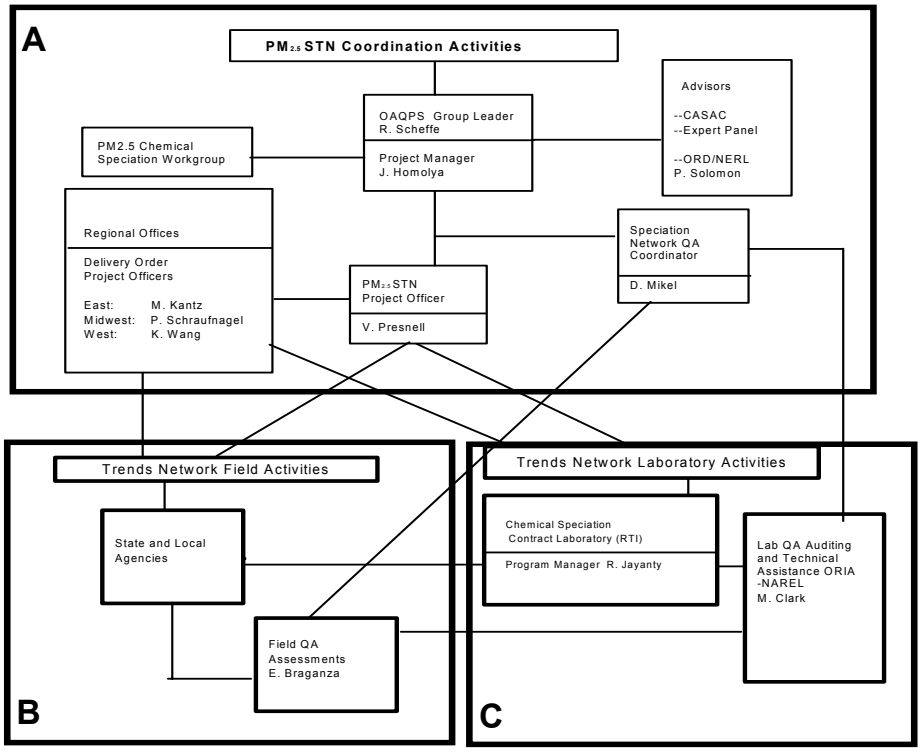


Figure 2. Management Structure

Agency	Type of Assessment	Agency Assessed	Frequency
NAREL	TSA, MSRs and PEs	RTI	Annually
R&IEL	TSAs, Performance Audits	State and local agencies	Annually*
OAQPS-EMAD	TSAs	RTI, NAREL, State and Local agencies, Regional offices and R&IEL	As needed by EMAD determination
OAQPS-EMAD	MSRs	RTI, NAREL, State and Local agencies, Regional offices and R&IEL	As needed by EMAD determination
Regional Offices	Network Reviews	State and local agencies	Once every 3 years

Table 1. Description of the Assessment

R&IEL – Las Vegas, Nevada: R&IEL will have QA oversight of the field operations. As such, the laboratory will perform TSAs and performance audits at state and local agency monitoring stations. At

this time, the schedule and extent of the TSAs and performance audits are not known. When the audits are completed, the QA reports will be forwarded to OAQPS for review. At some time in the future, State Agencies and the Regional offices will be encouraged to perform these functions.

Regional Offices: The EPA Regional Offices will provide Network Reviews of the STN on each agency within their region once every three years.

Since EPA is providing funding for this program, the QA requirements fall under the auspices of EPA Order 5360.1 July, 1998. In short, this EPA order states that all extramural activities funded by the EPA must have minimum requirements in place at the beginning and carried through the project. In order to fulfill this EPA Order, the EPA has put the following QA components in place.

EPA Method 1631 Quality Control for Sampling: Closing the Loop

William A. Telliard, Statistics and Analytical Services Branch, EPA Office of Water
Harry B. McCarty and Judy Schofield,
DynCorp Systems and Solutions, Science & Engineering Group

In June 1999, the EPA Office of Water developed, proposed, and promulgated Method 1631 for the determination of mercury in aqueous samples at levels as low as 0.2 ng/L (parts per trillion). The method employs oxidation, purge and trap, desorption, and cold-vapor atomic fluorescence spectrometry procedures to achieve the sensitivity needed to demonstrate compliance with the EPA water quality criteria for mercury. Because contamination can significantly affect sample results at these trace levels, the successful application of Method 1631 requires careful control of all sources of mercury contamination from the time that the sampling personnel begin to plan the project and continuing throughout sample collection, shipment, processing, and analysis. Such control requires up-front preparation, training of personnel, and closing the loop from the final results back to the sample collection process to evaluate the potential of the sampling procedures to contribute to the final analytical result and to determine when corrective actions are required.

Section 303 of the Clean Water Act (CWA) requires states to set a water quality standard for each body of water within its boundaries. A state water quality standard consists of a designated use or uses of a waterbody, the water quality criteria that are necessary to protect the designated use, and an antidegradation policy. The 1987 amendments to the CWA required states to adopt numeric criteria for toxic pollutants. These water quality criteria are designed to protect aquatic life, wildlife, and human health.

The U.S. Environmental Protection Agency (EPA) has established water quality criteria under the auspices of the National Toxics Rule (58 FR 60848) and the Stay of Federal Water Quality Criteria for Metals (60 FR 22228), and codified those criteria at 40 CFR 131.36. In addition to the water quality criteria published at 131.36, EPA has established water quality criteria in the Water Quality Guidance for the Great Lakes System at 40 CFR 132. The lowest water quality criterion for mercury is a criterion for the Great Lakes System for protection of wildlife of 1.3 ng/L. EPA developed Method 1631 to specifically address State needs for reliable measurement of mercury at water quality criteria levels.

Measurement of mercury by Method 1631 is accomplished by oxidizing the mercury in a sample with bromine monochloride (BrCl), using ammonium hydroxide and stannous chloride to convert Hg^{+2} to volatile Hg^0 (elemental mercury), purging the volatile Hg^0 from water onto a gold-coated sand trap, thermally desorbing the trap, and detecting the mercury by cold-vapor atomic fluorescence spectrometry (CVAFS). In developing this and other 1600-series trace metals methods, EPA Office of Water found that one of the greatest difficulties in measuring pollutants at these levels was precluding sample contamination during collection, transport, and analysis. Method 1631, therefore, requires contamination control through the use of "clean" techniques to collect and analyze the sample.

The philosophy behind contamination control and "clean" techniques is to reduce or eliminate contamination in order to produce a reliable results. The basis of this philosophy is given in Method 1669, the sampling guidance developed by EPA Office of Water to accompany Method 1631:

"The philosophy behind contamination control is to ensure that any object or substance that contacts the sample is nonmetallic and free from any material that may contain metals of concern."

This means that steps are taken to eliminate or reduce the presence of mercury in the sample bottles, sampling equipment, reagents, laboratory, labware, and laboratory air. Mercury also must be prevented from entering the sample at the sampling site. Mercury concentrations must be monitored in sampling and analytical equipment to ensure reliable measurements of mercury in environmental samples. To that end, Method 1631 specifies the collection and analysis of a series of quality control samples that are designed to monitor the sample collection processes for any spurious contributions of mercury to the final results. These QC samples are in addition to the traditional laboratory QC samples and include: field blanks, bottle blanks and sampler check blanks.

None of these QC samples are particularly new. In the mid-1970s, scientists identified the need to control contamination in order to make meaningful measurements of mercury and other metals at part-per-trillion levels in open ocean environments and in air samples at remote locations such as the South Pole, and they developed QC samples as a means to evaluate the potential for contamination throughout the sampling and analysis process.

In comparison to other methods, Method 1631 explicitly links reporting compliance monitoring results to acceptable results for the QC samples. The linkage makes communication and feedback among the laboratory, the sampling personnel, and the permittee a critical aspect of making meaningful measurements of mercury for compliance monitoring. The QC samples enable the parties to pinpoint contamination sources associated with different components of the sampling and analysis processes and to adjust the procedures in order to accurately characterize the mercury concentrations of environmental samples. These QC samples and the information that they provide regarding the contamination associated with the components of the sampling and analytical process are listed in Table 1. The method describes how these QC samples can be used to evaluate the quality of the sample results and make decisions about the utility of the sample results for monitoring compliance with the water quality criteria or other regulatory limits.

Table 1. Method 1631 QC Sample Description

QC Sample	Description	Frequency	Performance Criteria	Component of Sampling and Analytical Procedure Addressed
Reagent Blank	The concentration of mercury is determined in a solution containing all of the analytical reagents used in the analysis.	Each new batch of reagents	<25 pg	Analytical Reagents
Bottle Blank	After being cleaned, a representative set of sample bottles are checked for cleanliness.	1 per cleaning batch	<0.5 ng/L, or one-fifth of the Hg in associated sample(s), whichever is greater	Sample bottles
Sampler Check Blank	Sampler check blanks are generated in the laboratory or at the equipment cleaning facility by processing reagent water through the entire sampling system using the same procedures that are used in the field.	1 following each cleaning of the sampling equipment	<0.5 ng/L, or one-fifth of the Hg in associated sample(s), whichever is greater	Sampling equipment and procedure (reproduced in the laboratory)
Field Blank	Collected from the same site at the same time as the field samples.	10% from same site at same time	<0.5 ng/L, or one-fifth of the Hg in associated sample(s), whichever is greater	Entire sampling and analytical process, including sample handling, shipment, and storage

For example, if the results for the field blank are not within the performance criteria of the method, then the other QC samples can be evaluated to determine the source of the contamination. If the field blank is contaminated and the sampler check blank and bottle blanks are within the performance criteria of the method, that suggests that the sample is being contaminated during sample collection, sample handling or sample storage. In such a case, additional efforts (corrective actions) can focus on training the sample collection personnel to avoid contamination at the sampling site, and on a review of sampling equipment such as personal protective gear, gloves, or other likely sources of mercury contamination.

Conversely, if the laboratory-specific QC samples, such as the reagent blanks indicate contamination, and the field blanks and sampler check blanks provide no additional indications of contamination, then the corrective actions can focus on the laboratory environment and materials.

Taken in total, these QC samples facilitate reliable determination of mercury in the environment at the sub-part per trillion levels necessary to demonstrate compliance with water quality criteria. What sets this approach in Method 1631 apart from other EPA methods for mercury or other pollutants is tying the

results from QC samples covering the entire collection and analysis and establishing performance criteria for these QC samples that must be met before reporting the final results.

The alternative is a situation all too well known: A sampling program is designed and includes a series of QC samples. The samples are collected, shipped to the laboratory and analyzed. The results of QC samples such as field blanks suggest contamination of the samples. The samplers blame the laboratory, the laboratory blames the samplers, the data user is left with ambiguous results that add no value to the overall effort, and the regulatory authority is forced to decide compliance based on inadequate data. Method 1631 cannot prevent the finger pointing, but by explicitly tying field QC results to reporting compliance data, it closes the loop between field and laboratory activities.

References:

USEPA, March 2001, Method 1631, Revision C, Mercury in Water by Oxidation, Purge and Trap, and Cold Vapor Atomic Fluorescence Spectrometry, EPA-821-R-01-024.

USEPA, March 2001, Guidance for Implementation and Use of EPA Method 1631 for the Determination of Low-Level Mercury (40 CFR part 136), EPA 821-R-01-023.

USEPA, July 1996, Method 1669, Sampling Ambient Water for Determination of Metals at EPA Water Quality Criteria Levels.

Improving Data Confidence for EPA's Whole Effluent Toxicity Tests (WETT)

Carl Craig, Mark Carter, Shawn Kassner, Jeff Lowry
Environmental Resource Associates

Proficiency testing is an integral component of quality management for environmental chemistry laboratories. Laboratories utilize proficiency testing samples to assess general laboratory performance, to demonstrate capabilities, troubleshoot problems, and so forth. Clearly, with the range of proficiency testing programs in existence, environmental chemistry laboratories have many options for independent analytical data quality assessment. Whole Effluent Toxicity Testing, on the other hand, is at a distinct disadvantage. The endpoint analyses for these biological tests have no "true" value. Therefore, the laboratory must rely on historical, single laboratory references to evaluate performance. NPDES Permitees that rely on data from commercial laboratories are also at a disadvantage because there are few valid mechanisms to evaluate laboratory competence or capability. This presentation is intended to address the framework for a WETT proficiency testing program, and to discuss the issues associated with such a program.

Purpose: Reasons for performing Proficiency Testing (PT) as part of analytical data generation include the following[1-2]:

- % To test the ability of a laboratory to meet data quality requirements for specific environmental and/or regulatory programs.
- % To meet certification program, or contractual requirements.
- % To assess performance against peer or other currently certified and accredited laboratories.
- % To conduct self-assessment.

The primary goal of a superior PT program is the improvement of overall laboratory performance. This is a critical benchmark for all stakeholders to be able to quantitatively benchmark, and demonstrate improvement. One should always be mindful of the fact that the ancillary objectives of any PT program cannot overshadow this primary reason for conducting proficiency tests. Continuous process improvement should always guide the development of proficiency testing; otherwise the purpose for the program should be reviewed.

Background: Proficiency Testing programs are certainly not new to the environmental laboratory community. Starting in the early 1970s, the USEPA EMSL-Cincinnati laboratory managed performance evaluation (PE) studies to support the Agency's drinking and wastewater programs. The results of these single blind PE studies were used by state laboratory accreditation agencies as part of their process for accrediting environmental laboratories. In general, a laboratory needed to successfully participate in these studies to obtain or maintain state accreditation.

The USEPA Offices of Water and Enforcement jointly established the objectives and requirements of the Agency's proficiency testing programs along with the EMSL-Cincinnati laboratory. EMSL-Cincinnati then defined the technical procedures necessary to carry out the programs, performed or managed every aspect of the studies and provided 'self-oversight' to assure that program objects

were met for every study conducted. Because of EMSL-Cincinnati's impartiality and credibility, state accreditation agencies and participant laboratories alike relied on the Agency's proficiency studies for many years without any significant question.

Due to budgetary and personnel constraints, in 1994 the EPA formed an internal work group to evaluate the possibility of privatizing their PE programs. On July 17, 1996, the work group published a Federal Register notice describing the availability of a document titled the *Externalization of EPA'S Water Laboratory Performance Evaluation Programs* (USEPA Office of Water, EPA 800-D-96-001). In the notice, the Agency described four goals for privatizing its PE program.

> *The redesigned program will not result in any significant changes to existing EPA regulatory requirements or compliance monitoring programs.*
>
> *Authorities delegated to the states under the Clean Water Act and the Safe Drinking Water Act and related federal regulatory provisions also will not change in any substantive way as a result of the redesigned program.*
>
> *EPA and the states will receive all of the information needed to fulfill the requirements of regulatory, compliance monitoring, and laboratory certification programs under the redesigned program.*
>
> *In order to facilitate reporting, electronic methods of transmission utilizing standardized data formats will be developed and implemented to the maximum extent possible.*

USEPA was proceeding under the premise that the private sector could and would, with the proper standards and controls, offer proficiency testing studies of comparable quality to the Agency's own PE studies. In 1998, EPA published a final rule on the privatization of the PE program.

Concurrent with EPA's effort, the National Environmental Laboratory Accreditation Conference (NELAC) developed a similar set of requirements for the NELAC PT program. In July 1997, the NELAC Proficiency Testing (PT) Standards were first approved.

In October 1999, NIST's National Voluntary Laboratory Accreditation Program (NVLAP) accredited the first class of PT providers. Immediately, NIST NVLAP accreditation became a prerequisite for supplying PT samples to laboratories in both NELAC and non-NELAC states. The EPA/NIST program included the 279 analytes that historically were included in the EPA's Water Supply (WS), Water Pollution (WP) and DMR-QA PE studies.

Privatized PT programs have developed quickly over the past several years. In 1990, under the EPA, there were 4 WP and WS studies conducted annually. Data turn around was slow. Currently with 12 NVLAP accredited PT providers, there are more than 100 possible PT studies available to all chemistry laboratories. There are over 800 analytes included in a variety of matrices. Response time (the ability to initiate participation in a PT study) is immediate, with studies being conducted continuously. PT Providers are investigating mechanisms for reducing the turn-around-time of results and reports. The benefit of these improvements and the pressure to continuously improve and evolve private PT programs is obvious for the data users.

Proficiency Testing and Bioassay: Whole Effluent Toxicity Testing has a significant role in the USEPA's water quality management programs. The regulatory justification for conducting WET testing stems from the Clean Water Act (CWA) enacted in 1972 where Section 101(a)(3) states that "it is the national policy that the discharge of toxic pollutants in toxic amounts be prohibited."[4] The concept is simple; assess aqueous toxicity not through chemical identification and quantitation, but by observing the effect of the effluent on organisms.

Whole Effluent Toxicity Testing programs have been in place since the late 1940's and early 1950's[3]. Whole Effluent Toxicity is officially defined as "the aggregate toxic effect of an effluent measured directly by an aquatic test"[4]. There are thousands of NPDES permittees with WET requirements written into their discharge permits. We estimate that half of all major permits are written with WET requirements. Placing this in perspective, WET requirements on NPDES permits are similar in number to some important chemistry parameters such as pH, demand or ammonia.

In 2000, the EPA conducted their final PE study for the Whole Effluent Toxicity (WET) program. This was the DMR-QA 20 study. Environmental Resource Associates conducted a cursory investigation in 2001, to understand toxicity laboratory satisfaction with the DMR WET program and found that the studies did not seem to be meeting the intent of a PT program as highlighted above. The program was perceived as required without providing benefit to the laboratory. Certainly the WET DMR-QA studies were not designed to provide rapid data feedback to laboratories and permittees.

It is also clear that there were benefits derived from the USEPA DMR-QA WET program. On one occasion we spoke to a laboratory manager who failed a recent DMR-QA WET study. Upon evaluation of the data the laboratory found that the organisms they were using were more mature and robust than would be allowable, and by modifying their test organisms were able to be more consistent with other accredited laboratories. This is clearly one of the objectives of a PT program. The disturbing summary of this discussion was that the perception of the laboratory manager was that DMR-QA for the WET study was not helpful for the laboratory. His laboratory was trying to use "compliant" yet robust organisms for their WET tests. The goal was to minimize impact of the effluent on the organism so the discharges would not be found to be toxic.

***Future of WET Proficiency Testing*:** Chemistry parameters have been at the center of government and now private PT programs for years. In the past 4 years chemistry Proficiency Testing programs have expanded significantly. Many of the improvements, changes and expansions are customer initiated. As a result, these programs have become focused on meeting the requirements described above for PT programs and are not used (and perceived) as only useful for compliance.

For example, chemistry laboratories may now acquire single-blind PT standards, with near limitless combinations of analytes to evaluate everything from new instrumentation to analyst proficiency. Laboratories can use PT programs to assist in troubleshooting laboratory data quality problems on a year-round basis. The results of chemistry PT programs are not always reported to regulatory agencies, prompting laboratories to use the programs as a proactive approach to improvement or problem solving. This is possible in part because so many laboratories find value in participating in these studies, and do so frequently.

Advances in computer systems and data collection have also increased the value of the private PT programs. For example, participants now have access to more detailed data analysis to determine exactly how well they performed within their peer group. This allows a more complete evaluation of performance over the binary pass-fail systems.

Development of a comprehensive WET Proficiency Testing program will begin with simple technical improvements. The first of these technical improvements were begun with the USEPA's Variability Study. Reference toxicants were prepared as ready to dilute concentrates, making the preparation of the simulated effluent very simple. Previously, all reference toxicants were supplied to laboratories as dry materials. From this the laboratory had to follow meticulous instructions to prepare the stocks and solutions correctly. This introduces some variability and is a source of

possible error. DMR-QA 22 will include the use of reference toxicant concentrates that are diluted directly to the 100% effluent concentration. This should have considerable impact in simplifying the preparation of the effluents for the 19 tests.

Secondarily, the summary of results will be available more expeditiously. Participants will be able to see preliminary information soon after the study close. This timely feedback affords labs the possibility of initiating corrective action investigations promptly. Responsiveness of this type in a PT program allows the laboratories to determine the temporal nature of problems with laboratory performance. When data is supplied months after the end of a study, it can be very difficult to determine if the problem is periodic, transient or a singular event.

***Challenges*:** Additional, and more frequent WET PT studies should also begin to occur. This will allow laboratories to evaluate their own improvements, corrective actions, and so forth, without the data having to be evaluated by a regulatory authority. Conducting regular studies with enough participating laboratories will clearly be a challenge, as there are far fewer laboratories conducting toxicity testing than analyzing chemistry parameters. However, the availability of "routine" PT studies has made chemistry PT studies more accepted and useful. We believe this will occur for WET as well.

Well-characterized toxicants are currently limited to a few well-known chemicals. Studies to develop endpoints for chemicals on a variety of organisms are very expensive. Yet, to make routine studies available, a library of toxicants and their impact on the classes of organisms used in aquatic bioassay needs to be developed. There are many sources of good information for past toxicity studies. These reside within the USEPA, EPA Regional Laboratories, States and in literature provided by industry groups such as the Society of Environmental Toxicology and Chemistry (SETAC). It will now be the burden of PT providers to review, collate and use this information to develop additional reliable toxicants.

Summary: Proficiency testing should be a successfully integrated part of all laboratory quality programs. The success and usefulness of PT studies originates with a clear understanding of the goals of the program. The goals should focus on the overall objective of improving data quality. In comparison, if the primary goal of a PT program is compliance assessment, then the primary goal for setting speed limits must be to give speeding tickets. There must be a viable and understood benefit to the participant laboratories in order to realize the full potential of any PT program.

For Whole Effluent Toxicity, developing a private PT program that is useful to the laboratories and their customers is possible. Technical improvements, together with developing more frequent and timely programs that support confidence in laboratory data while improving quality and defensibility is a realistic and necessary goal in support of the WET program.

References:

1. Taylor, John K. *Quality Assurance of Chemical Measurements:* 1987, CRC Press, Chapter 23.

2. *Proficiency Testing,* National Environmental Laboratory Accreditation Conference, Rev. 16, May 25, 2001.

3. *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*; Grothe, D.R., Dickson, K.L., Reed-Judkins, D.K., Eds. 1996, SETAC

Press, Chapter 2.

4. *Final Report: Interlaboratory Variability Study of EPA Short-term Chronic and Acute Effluent Toxicity Test Methods, Vol.1,* United Stated Environmental Protection Agency, Office of Water, EPA 821-B-01-004, September 2001, pp. 1.

One Fish, Two Fish, We QC Fish: Controlling Data Quality Among More than 50 Organizations over a Four-Year Period

Lynn Riddick, DynCorp Environmental
Cynthia Simbanin, U.S. Environmental Protection Agency

EPA is conducting a National Study of Chemical Residues in Lake Fish Tissue. The study involves five analytical laboratories, multiple sampling teams from each of the 48 participating states, several tribes, all 10 EPA Regions and several EPA program offices, with input from other federal agencies. To fulfill study objectives, state and tribal sampling teams are voluntarily collecting predator and bottom-dwelling fish from approximately 500 randomly selected lakes over a 4-year period. The fish will be analyzed for more than 300 pollutants. The long-term nature of the study, combined with the large number of participants, created several QA challenges: 1) controlling variability among sampling activities performed by different sampling teams from more than 50 organizations over a 4-year period; 2) controlling variability in lab processes over a 4-year period; 3) generating results that will meet the primary study objectives for use by OW statisticians; 4) generating results that will meet the undefined needs of 50+ participating organizations; and 5) devising a system for evaluating and defining data quality and for reporting data quality assessments concurrently with the data to ensure that assessment efforts are streamlined and that assessments are consistent among organizations. This paper describes the QA program employed for the study and presents an interim assessment of the program's effectiveness.

Introduction

The *National Study of Chemical Residues in Lake Fish Tissue* is a four-year, multi-million dollar effort led by EPA's Office of Water (OW) with extensive participation by 48 states, several tribes, each of the EPA Regional offices, and several EPA program offices. The primary objective is to estimate the national distribution of selected persistent, bioaccumulative, and toxic (PBT) pollutants in fish tissue from lakes and reservoirs of the continental U.S. However, results from the study will be useful to OW and other participants for a variety of other purposes, such as providing information to 1) fulfill objectives of the Clean Water Action Plan (CWAP), 2) support the EPA's PBT initiative, 3) answer important questions concerning the regional occurrence of fish tissue contamination, and 4) suggest specific areas in need of further study. Given the variety of potential uses for the data, the broad number of participants, the four-year duration, and the broad geographic range of the study, it was clear that a strong quality assurance (QA) program would be needed. Recognizing this, OW managers included quality management activities throughout every phase of the study, beginning with the earliest planning phases. The result was a comprehensive QA program that addressed all aspects of study planning and implementation and that provides an effective means for real-time assessment and, where needed, improvement while the study is underway.

QA Program Elements
Planning

Collaborative Study Design: OW tasked a team of statisticians, biologists, and chemists from its Office of Science and Technology (OST) to work closely with experts from the Environmental Monitoring Assessment Program (EMAP) within EPA's Office of Research and Development (ORD) and with chemists in EPA's Office of Prevention, Pesticides, and Toxic Substances (OPPTS) to design the study. Together, the team developed a basic design framework and a draft list of possible target pollutants. The framework and proposed pollutant list were documented in a draft Study Design Document, along with an explanation of the team's rationale for the proposed design and target analyte list, issues considered by the team, and specific areas that the team felt should be further considered by a broader list of experts.

In October 1998, OW convened a two-day, invitation-only workshop to obtain peer input on the draft plan. More than 50 experts from federal, state, and tribal organizations (including EPA, NOAA, USGS, and state environmental, wildlife, and fisheries management agencies) were invited to participate. The draft study design document and a peer review charge were distributed in advance so that participants could arrive prepared with specific comments, questions, and concerns. After describing the study objectives, resource limitations, draft design, and specific areas of concern, workshop participants were split into workgroups to consider four study design issues: 1) sampling design or approach; 2) pollutants of concern; 3) sampling methods; and 4) data management. The workshop concluded with workgroup presentations of their findings and recommendations for further consideration by OW study managers.

Use of the collaborative study design approach allowed OW to develop a study design that reflected lessons already learned by other organizations. At the same time, the design afforded an opportunity to build interest and gather support from the states, tribes, and Regions, on whom OW depended to collect study samples.

Use of Workgroup to assist in Method Identification: In order to identify methods or techniques that would best meet OW's needs for the study, OW invited a group of recognized experts in the field of fish tissue analysis to participate in an Analytical Methods Workgroup. The workgroup was asked to assist in reviewing method development strategies and draft methods. Using guidelines suggested during the study design workshop, the workgroup sought to identify techniques that 1) minimized method development and validation costs; 2) yielded the lowest possible detection and quantification limits, and, where possible; 3) avoided the use of expensive or highly novel analysis techniques that could increase analytical costs. This process identified two existing methods that met study needs with no further testing, and several methods that required slight modification or testing to add target pollutants or optimize performance in tissue.

Selection of methods with QC elements that support quality objectives of the study: All of the methods used in the study, including those that were modified to meet study objectives, detail a comprehensive suite of laboratory QC elements needed to control and define the quality of results produced by each lab. These elements include: 1) required use of pure and traceable reference standards, 2) procedures for verifying that required detection and quantification levels are achievable by the laboratory, 3) procedures for demonstrating that the instrument is properly calibrated prior to and throughout sample analysis, 4) procedures for preparing, analyzing, and evaluating laboratory QC samples before analysis and during each shift to demonstrate the laboratory's ability to obtain precise and accurate results with the method, 5) use of either matrix spike samples or isotopically labeled standards to quantify recoveries of target analytes from tissue samples, and 6) required analysis of blanks to demonstrate the absence of contamination.

Use of Approved Quality Assurance Project Plans (QAPPs) for Sampling and Analysis Activities: Two QAPPs were developed and approved by EPA to support this study. The *Quality Assurance Project Plan for Sample Collection Activities for a National Study of Chemical Residues in Lake Fish Tissue* (May 2000) establishes data quality goals for all sample collection and handling activities and describes the QA/QC techniques employed by field teams and by the field support contractor to support those goals. The *Quality Assurance Project Plan for Analytical Control and Assessment Activities in the National Study of Chemical Residues in Lake Fish Tissue* (September 2000) establishes measurement quality objectives (MQOs) for laboratory data generated during the study and describes QA/QC techniques that employed by laboratory and sample control contractor staff to ensure these MQOs are met.

Field Orientation/Training Program: Because the study design relied on a large number of state, tribal, and Regional sampling teams, the Office of Water conducted regional field orientation and training programs to ensure that personnel responsible for sampling activities within each organization understood the study objectives, were familiar with customized-paperwork developed to document sample collection activities, and were prepared to collect, document, and ship samples in accordance with the study design and the sample collection QAPP.

Implementation

Distribution of Study-Specific Sample Documentation and Sampling Kits: The study design calls for collection of fish samples by multiple teams from participating states, tribes, and EPA Regions. To date, more than 120 samplers representing 56 organizations have participated in the study. To ensure samples will be consistently documented by such a large and diverse group, several forms were custom-designed for the study. These forms include a *Field Record Form* to document information about each lake sampled and individual specimens collected from the lake, a *Sample Identification Label* to accompany and identify each fish specimen, and *Chain-of-Custody* documentation. These forms are included in *custom-made sampling kits* prepared and distributed by EPA's sample control contractor. The kits also contain contaminant-free materials needed to store each specimen, a reference instruction sheet with contact phone numbers, and pre-completed forms needed to ship the specimens to the Sample Prep Laboratory for homogenization and compositing. Finally, sample *Traffic Reports* were created for use by the Sample Prep Lab to document each homogenized composite aliquot sent to either an Analysis Lab or to the Sample Repository for long-term storage.

Use of a Single Sample Prep Laboratory to Homogenize and Composite Samples: Many of the pollutants targeted in the study are being measured in the part per trillion or part per quadrillion range using state-of-the-art measurement techniques. For example, the quantification limit for mercury in tissue for this study is 2 ng/g, dioxin is being quantified at 0.1 ng/kg, and individual PCB congeners are being quantified at levels as low as 1 ng/kg. (Detection limits are even lower than these figures.) With monitoring levels this low for ubiquitous pollutants, it is critical to ensure that levels reported in samples reflect the true concentration of pollutants in samples and are not the result of contamination. Therefore, all sample processing activities (i.e., filleting, grinding, homogenizing, compositing, and aliquotting) are performed in a strictly controlled, clean laboratory and are associated with QC samples that will capture any problems with the sample prep processes.

Use of a Single Laboratory to Analyze a Given Pollutant throughout the Duration of the Study: As noted above, the study is statistically designed to determine the national distribution of pollutant residues in lake fish tissue. In any statistical analysis, lower measurement error translates a higher level of confidence in final results. Nearly all of the QC measures described in this paper were designed to reduce sources of error. Use of a single laboratory to make all

measurements for a given pollutant provided a rare opportunity to eliminate one source of error--interlaboratory variability.

Prequalification of Laboratories: Prior to analyzing any samples collected in the study, each Analytical Laboratory was required to submit acceptable method detection limit (MDL) and initial precision and recovery (IPR) study results generated in appropriate reference tissue matrix using the analytical method they would be using in the National Fish Tissue Study. MDL studies, which involve analysis of seven low-level (i.e., in the detection limit range) replicate samples, were to be conducted in accordance with the procedures given at 40 CFR 136, Appendix B. IPR studies, which involve preparing and analyzing four replicate reference standards spiked within the measurement range, were to be performed in accordance with the procedures given in each method.

Strong Communication Network: Routine contact with project staff and project participants is an integral aspect of the study design that has significantly contributed to the overall quality of data gathered in the study. The communication network employed in this study varies both in frequency (i.e., daily, weekly, monthly, annually) and medium (i.e., meetings, phone, email, fax), according to need. Highlights of this communication network include:

- **Daily monitoring of sampling and laboratory activities**: OW's contractor teams have been tasked with daily coordination and monitoring of sample collection, shipment and analysis activities. This monitoring has prevented unnecessary thawing of samples when shipping delays occurred and allowed OW to mitigate the impacts of deviations from the study design, thereby ensuring that limited study resources are used appropriately.
- **Monthly project meetings:** Each month, the study manager holds a meeting to discuss study status, schedules, and issues with other OW staff responsible for managing laboratory and data review activities and with the team of contractors responsible for daily tracking of activities. Depending on project activities, additional staff are brought into these meetings to facilitate planning and resolution of issues.
- **Feedback to Study Participants:** The OW Study Manager regularly communicates with study participants concerning the study status, issues, and concerns. Broad issues that affect all participants are disseminated via email. Examples include dissemination of progress reports, clarifications concerning the amount of dry ice needed for shipping, and requests to halt following the September 11, 2001 attacks. Specific concerns are discussed via telephone.
- **Annual reporting of results**: Analytical results, and associated data quality assessments, are being reported back to each state, tribe, and region on an annual basis so that these organizations can evaluate their results and resolve any questions about their data prior to public release. Public release of study results is delayed for six months to accommodate such reviews.

Implementation of standardized data format: All data generated during the study are being compiled in a centralized, custom-developed database to ensure that results are reported to users consistently. The database allows for: 1) eventual upload of results to the national STORET database system, 2) statistical manipulation of results, 3) export of results to user-friendly formats such as Excel spreadsheets, and 4) consistency in data format and nomenclature across laboratories and over time.

Assessment

Three Levels of Data Quality Assessment and Application of Standard Data Qualifiers: All analytical data generated in the study are subjected to three levels of review. First and as noted above, a pre-qualification review was performed on data submitted by each laboratory to demonstrate that the labs were qualified to prepare and/or analyze tissue samples collected during the study. Second, each submission of sample results is carefully scrutinized to verify that the samples were analyzed as directed and that supporting QC results demonstrated the quality of results generated. In evaluating these submissions, data reviewers employ a suite of standardized data qualifiers and abbreviated qualifier codes to consistently and accurately document the quality of all data generated so that both the primary data users (statisticians) in EPA Headquarters and secondary data users within Regions, states, tribes, and other organizations can make informed decisions regarding data use. A third level of data review is performed at the conclusion of each year and, ultimately, at the conclusion of the study, to determine if overall data quality supported study objectives.

Documentation of Data Quality in Annual QA Reports: With more than 50 different organizations interested in using data from this study, it is easy to imagine an unnecessary duplication of resources among organizations assessing data quality. To avoid this, OW thoroughly documented the procedures it was using to review the data, flag the data, and define data quality in a *Quality Assurance Report for the National Study of Chemical Residues in Lake Fish Tissue: Year 1 Analytical Data*. OW also documented all of its data quality findings in the report, and disseminated the reports to each study participant, along with the reviewed, flagged, and qualified data.

Improvement

The study QA program recognizes that unanticipated challenges can arise and includes mechanisms to take corrective actions on specific situations and make programmatic changes that can minimize the potential for future problems of the same nature. For example, sample shipping instructions had to be modified to direct samplers to declare their shipments as having $100 value after some Customs agents delayed shipping when they questioned how 75 lb coolers could be valued at less than $10. A more dramatic example occurred when semi-volatile analysis activities started. Although preliminary testing suggested that a slightly modified version of Method 1625 would be capable of handling tissue matrices, the lipid contents encountered in the first batch of samples analyzed resulted in excessive interferences and the need for repeated reanalysis. Rather than allow such problems to result in excessive delays and reanalysis costs throughout the study, additional method modifications were developed and tested before initiating further analysis of study samples.

Is it Working?

The first year of the study has been completed, and analytical activities for Year 2 are well underway. Data quality assessments from Year 1 indicated that:

- The Year 1 data set exceeded predefined study MQOs for precision, bias, sensitivity, and analytical completeness.
- 99.9% of the more than 100,000 field results gathered in Year 1 met all instrument calibration requirements.
- 98. 8% of the Year 1 field sample results were not affected by blank contamination of any kind during the study.
- 99.7% of the Year 1 field sample results had no QC problems that would suggest sample matrix interferences.

- 99.5% of the Year 1 results had no QC problems that would suggest laboratory performance problems.
- 99.8% of the more than 100,000 sample results were determined within analytical holding times (even when re-analysis was required).

Quality Control in Whole Effluent Toxicity Test Methods: Lessons from EPA's Interlaboratory Variability Study

William A. Telliard and Marion Kelly
Statistics and Analytical Support Branch, EPA Office of Water
Robert N. Brent and Harry B. McCarty
DynCorp, Science & Engineering Group

Whole effluent toxicity (WET) test methods are biological test procedures that measure the effects of pollutants on living organisms. These tests monitor the survival, reproduction, and growth of fish, invertebrates, and algae when exposed to pollutants. In 1999-2000, EPA conducted a large-scale interlaboratory variability study of 12 individual test methods and collected test data for over 700 samples analyzed in 56 laboratories.

WET test data are derived from an extremely large number of manual measurements and observations. For instance, up to 960 individual biological measurements and observations are required to derive a single test result for the analysis of a single effluent sample. All of these measurements are made visually, recorded manually, and manually entered into statistical software to generate the test result. The large number of measurements, the manual nature of data recording, and the number of laboratories reporting data in this study posed unique quality control challenges. To address these issues, EPA implemented a number of specific quality control measures including laboratory prequalification, adherence to promulgated test methods and study-specific standard operating procedures, test-specific QC criteria, data reporting standards, and independent data review and result recalculation.

The goal of the interlaboratory study was to characterize the performance of the WET test methods and assess the adequacy of these methods for use in a national monitoring program. The calculation of method QC performance measures from this study revealed that these methods exhibit interlaboratory variability comparable to chemical methods approved for national monitoring and below previous estimates for WET methods.

In addition to this finding, the WET Variability Study also highlighted other QA/QC insights including the benefits of hardcopy and standard electronic reporting formats for large studies, the benefits of independently recalculating test results in validation studies, and the importance of clearly defining test procedure flexibility.

Introduction

The Clean Water Act (CWA) was enacted in 1972 with the objective of "restoring the chemical, physical, and biological integrity of the Nation's waters." Along with other goals, CWA section 101(a)(3) states that "it is the national policy that the discharge of toxic pollutants in toxic amounts be prohibited." The U.S. Environmental Protection Agency (EPA) was tasked with meeting the CWA objectives and has pursued this goal through implementation of water quality standards and effluent permitting, monitoring, and compliance programs. Because the CWA mandate specifically prohibits "toxic pollutants in toxic amounts," the use of biological test methods are an integral part of these programs. Only biological test systems that measure the effects of pollutants on living organisms can directly assess toxicity. For this reason, EPA has integrated whole effluent toxicity (WET) testing into its toxics control strategy (USEPA, 1991).

Whole effluent toxicity (WET) is defined as "the aggregate toxic effect of an effluent measured directly by an aquatic toxicity test" [54 FR 23686; June 2, 1989]. WET tests are biological test procedures that expose living aquatic organisms (fish, invertebrates, or algae) to a range of effluent concentrations under controlled laboratory conditions. The organisms are exposed to the effluent for 24 hours (acute tests) to 7 days or more (short-term chronic or chronic tests). At the end of the exposure period, the survival, growth, and/or reproduction of the organisms is measured in each effluent concentration and a control treatment. Toxicity of the effluent is determined by statistically comparing measured responses between the control and various effluent concentrations.

Over the past several decades, EPA has developed a number of different WET test methods for measuring the acute and chronic toxicity of effluents to various species of fish, invertebrates, and algae. In 1995, EPA approved 17 WET test methods for use in monitoring compliance with discharge permits prohibiting toxicity. To resolve judicial challenges to this action, EPA initiated an interlaboratory variability study (the WET Variability Study) to validate and ratify 12 of these methods in 1999.

The WET Variability Study

From September 1999 through April 2000, EPA conducted the largest interlaboratory study of WET methods undertaken to date. This study assessed 12 different WET test methods, which included a variety of acute, chronic, freshwater, and marine methods. For each method, four different sample types were evaluated in up to 35 laboratories. A total of 56 laboratories were involved in the study, each evaluating an average of 3 different methods and analyzing an average of 13 samples. In total, over 700 individual WET tests were conducted in the WET Variability Study, and over 100,000 test organisms were used to conduct these tests.

The purpose of the WET Variability Study was to characterize the performance of the WET test methods through the evaluation of the following method QC performance measures:

- **Interlaboratory Variability** - the variability of test results when different laboratories test the same sample
- **Successful Test Completion Rate** - the rate at which qualified laboratories can successfully complete WET tests
- **False Positive Rate** - the rate at which WET tests indicate toxicity is present when measuring non-toxic samples

Based on the evaluation of these method QC performance measures, EPA assessed whether the existing approach to quality control in the WET test methods is effective and adequate for use in a national monitoring program.

Quality Assurance / Quality Control Elements

To accurately characterize the WET method QC performance measures, it was necessary to maintain strict quality control throughout all aspects of the study. The large scale of the WET Variability Study and the nature of the biological test data collected posed several unique quality control challenges that were addressed in this study. First, WET tests require an extremely large number of manual measurements and observations. Up to 960 individual biological measurements and recorded observations are required to derive a single test result for a single analyzed sample. In addition to those biological measurements, up to 378 physical and chemical measurements (e.g., temperature, dissolved oxygen, pH, etc.) may be recorded to monitor the conditions during each test. In total, over 300,000 individual biological, physical, and chemical measurements were conducted and reported in the WET Variability Study. Many of these measurements are performed visually (e.g., counting the number of offspring), recorded manually, and manually entered into statistical software to generate test results. The large number of measurements and the manual nature of data recording increases the chance of errors in reported data and complicates the quality control process. Secondly, these large numbers of

measurements were conducted and data were reported from 56 different laboratories, further complicating the quality control process. The following specific quality assurance and quality control elements were implemented in the WET Variability Study to address these issues:

- **Laboratory Prequalification** - To ensure that representative data of appropriate quality were used to evaluate WET test performance measures in this study, laboratories were required to submit specific prequalification documentation to demonstrate that they possessed the capacity and capabilities, experience and proficiency, and quality assurance and quality control systems necessary to meet the needs of the study. EPA evaluated all submitted documentation and selected participant laboratories that met established prequalification criteria.
- **Standard Operating Procedures** - Participant laboratories were required to analyze samples according to the promulgated WET test method manuals and specific instructions that EPA provided in the form of study-specific standard operating procedures (SOPs).
- **Test-Specific QC Criteria** - One of the goals of the WET Variability Study was to determine if the existing approach to quality control in the WET test methods is effective and adequate for use in a national monitoring program. For this reason, the standard quality control measures in WET testing, including test acceptability criteria, reference toxicant testing, and test condition monitoring were maintained in this study. Test acceptability criteria are minimum requirements for the performance of test organisms under blank, or negative control, conditions. Reference toxicant testing involves the routine analysis of a known toxic substance (a positive control) to provide ongoing control of test variability. Test condition monitoring provides control of the physical and chemical conditions under which tests are conducted.
- **Data Reporting Standards** - Participant laboratories were required to submit all hardcopy benchsheets of recorded measurements. In addition, laboratories were required to submit all data electronically in pre-designed standard reporting templates (created in Microsoft Excel®). These electronic reporting templates were designed to be user-friendly, capture all pertinent test information, and perform automated reviews of QC criteria.
- **Independent Data Review and Result Recalculation** - EPA independently reviewed all test-specific QC criteria to ensure that tests were conducted properly and met acceptability criteria. EPA also performed a quality control review of data entry by cross-referencing all biological data reported electronically with hardcopy laboratory benchsheets. EPA then independently re-calculated all test results using the reviewed electronic data. Re-calculated results were compared to laboratory-reported results to identify and resolve any inconsistencies. These steps ensured that all test results were consistently and accurately calculated.

Study Results

Following test review, results were compiled and method performance characteristics (interlaboratory variability, successful test completion rate, and false positive rate) were calculated for each WET test method. Table 1 displays summarized results from the WET Variability Study. Successful test completion rates were greater than 90% for all WET test methods except the *Ceriodaphnia* chronic (82%) and *Selenastrum* chronic (63.6%) test methods. False positive rates were less than 5% for all WET test methods. Interlaboratory variability was described by the coefficient of variation (CV) calculated for point estimates. Interlaboratory CVs of LC50s (median lethal effect concentrations) ranged from 20.0% to 38.5% for acute test methods. Interlaboratory CVs of IC25s (25% inhibition concentrations) ranged from 10.5% to 43.8% for chronic test methods. The variability of WET test methods measured in this study was comparable to that of chemical methods approved by EPA for use in national monitoring

programs. The variability measured in this study also was lower than measured for these methods in previous studies and reported at the time of method promulgation. Interlaboratory CVs measured in the WET Variability Study were 4% to 34% lower than average values cited for the same methods at the time of promulgation.

Table 1. Summarized test results from EPA's WET Variability Study

Test method	Successful test completion rate (%)	False positive rate[a] (%)	Interlaboratory Precision (%CV)[b]
Ceriodaphnia acute	95.2	0.00	29.0
Ceriodaphnia chronic	82.0	3.70	35.0
Fathead acute	100	0.00	20.0
Fathead chronic	98.0	4.35	20.9
Selenastrum chronic (with EDTA)[c]	63.6	0.00	34.3
Mysidopsis chronic	97.7[d]	0.00	41.3
Sheepshead acute	100	0.00	26.0
Sheepshead chronic	100	0.00	10.5
Silverside acute	94.4	0.00	38.5
Silverside chronic	100	0.00	43.8
Champia chronic[e]	ND	ND	ND[f]
Holmesimysis acute[e]	ND	ND	ND

[a] False positive rates reported for each method represent the higher of false positive rates observed for hypothesis testing results or point estimates.

[b] Coefficients of variation (CVs) reported for each method represent the CV of LC50 values for acute test methods and IC25 values for chronic test methods. CVs reported are based on total variance and averaged across sample types.

[c] The *Selenastrum* chronic test method was conducted with and without ethylenediaminetetraacetic acid (EDTA) as a component of the nutrients added to test and control treatments.

[d] Successful test completion for the optional fecundity endpoint was 50%.

[e] ND = not determined. Due to insufficient laboratory support, interlaboratory data were not obtained for the *Champia* chronic and *Holmesimysis* acute test methods.

[f] While interlaboratory test data were not obtained for the *Champia* chronic method, intralaboratory data was obtained from the referee laboratory. Intralaboratory CVs were 27.6%, 49.7%, and 50.0% for reference toxicant, receiving water, and effluent sample types, respectively.

QA/QC Lessons Learned from the Study

From this large-scale study, the following lessons were learned regarding QA/QC:

- **Adequacy of Test Specific QC Criteria** - The results of the study demonstrated the effectiveness of the existing WET QC measures. Using these test-specific QC criteria, most methods were able to achieve acceptable variability, high test completion rates, and low false positive rates. This conclusion is not to say, however, that WET test performance cannot be improved. In fact, EPA has proposed additional QC measures to control within-test variability. Controlling within-test variability would bring greater consistency to the level of test sensitivity achieved across laboratories and individual tests.

- **Benefits of Hardcopy and Standard Electronic Reporting Formats** - The study requirement to submit hardcopy laboratory benchsheets and electronic data in standard formats was extremely helpful in the review of data from such a large study. The cross-reference of data between the two sources allowed errors to be identified and corrected. Standard electronic formats also allowed automated review of QC criteria, direct import of data into statistical analysis software, and compact and accessible data archival. The large number of data sets generated by this study and the size of data sets generated by WET tests increased the usefulness of these quality control steps.
- **Benefits of Independently Recalculating Test Results** - Due to the importance placed on method validation studies, EPA felt that it was necessary to independently recalculate all test results in the WET Variability Study. This step ensured consistency in the generation of the data set. A comparison of laboratory-reported and independently-recalculated test results from this study revealed inconsistencies in the calculation or reporting of 54% of tests. While these inconsistencies were generally minor (with 63% resulting in <1% difference in the reported result), all of these inconsistencies were able to be identified through independent recalculation, and thus the consistency of the data set from this study was not affected.
- **Clarification of Test Procedure Flexibility** - Despite the required use of promulgated test procedures and study-specific SOPs, WET tests occasionally do not meet all of the specified test conditions. This is often merely a product of the large number of test condition measurements (e.g., up to 378 per test) and narrow recommended ranges (e.g., ±1°C for temperature). For this reason, the WET method manuals have allowed flexibility in the conduct and review of WET test data. This study has highlighted the need to clearly define this allowable flexibility in the method and clearly distinguish between mandatory and recommended test procedures or conditions.

Analysis of the OAQPS Policy on the Secondary Use of Existing Data

R. Wright, K. Morgan, I. Beaty
RTI International

OVERVIEW OF OAQPS SECONDARY DATA POLICY

Status of OAQPS Secondary Data Policy

EPA Order 5360.1 A2, concerning the Agency-wide quality system, applies to environmental programs that make secondary use of existing data (secondary data). The OAQPS secondary data policy is given in its Quality Management Plan, which received Agency approval in September 2001: "It is the policy of OAQPS that within the constraints of available resources, QA activities associated with secondary data shall be conducted to assure that data will be adequate and sufficient for their planned secondary use." In general, there should be sufficient information about data quality to determine whether they can be used to support a particular environmental decision.

In order to assist in the implementation of this policy, OAQPS issued a Task Order for an independent assessment of this policy. RTI International (RTI) was awarded this Task Order, and has completed this work as a result of that Task Order.

Tiered QA Approach to Secondary Use of Existing Data

In accordance with OAQPS policy, available resources for important projects are maximized through the use of four QA categories for projects. The number of QA elements required for each category is reduced as one proceeds from Category I (most stringent) to Category IV (least stringent)[1].

Category I:	requires all A, B, C, and D elements of a QA project plan
Category II:	requires all elements of a Category I project plan, except A9 (Special Training Requirements/Certification)
Category III:	requires most of the A and B elements, the C elements, and the D elements
Category IV:	requires A1, A6, A7, B1, B4, B5, but none of the C or D elements are required

All environmental data operations must be covered by an approved QA Project Plan or equivalent planning documentation prior to the start of the project. Implementation of the OAQPS policy is more likely to succeed if concise guidance is provided along with useful tools and templates for those who use existing data.

Requirements and Guidance for QA Project Plans Involving Secondary Use of Existing Data

EPA Requirements for Quality Assurance Project Plans (EPA QA/R-5) includes the following specifications for non-direct measurements, a.k.a. secondary data, under Element B9: (1) identify any types of data needed for project implementation or decision-making that are obtained from non-measurement sources such as computer databases, programs, literature files, and historical databases; (2) describe the intended use of the data; and (3) define the acceptance criteria for the use of such data in the project and discuss any limitations on the use of the data.

EPA Guidance for Quality Assurance Project Plans (EPA QA/G-5) states that Element B9 should clearly identify the intended sources of previously collected data. Information that is non-representative and

[1] From Table 2-1 of the 2001 *OAQPS Quality Management Plan*.

possibly biased and is used uncritically may lead to decision errors. The care and skepticism applied to the generation of new data are also appropriate to the use of previously compiled data.

SUMMARY OF CURRENT PRACTICES AND VULNERABILITIES RELATING TO OAQPS SECONDARY DATA POLICY

As part of this project, RTI conducted a telephone survey with OAQPS staff to assess current practices and vulnerabilities regarding the policy. One important fact that was mentioned was that the White House Office of Management and Budget (OMB) recently issued guidelines on information quality (IQ) disseminated by federal agencies. The agencies are directed to adopt a basic standard of quality (including objectivity, utility, and integrity) as a performance goal and to incorporate information quality criteria into information dissemination practices. They are to adopt specific standards of quality that are appropriate for the various categories of information they disseminate. Where appropriate, data supporting disseminated information should have full, acccurate, and transparent documentation. Error sources affecting data quality should be identified and disclosed to users. Implementation of the OAQPS secondary data policy and the four-tiered project category approach may become important in the OAQPS response to the OMB guidelines.

Current Practices related to OAQPS Secondary Data Policy

Secondary uses of existing data. The Aerometric Information Retrieval System (AIRS) is a major source of primary data in OAQPS. It is considered primary data because they are being used by OAQPS for the specific uses for which they are collected. Outside of AIRS, OAQPS uses more existing data than primary data. Fiscal limitations sometimes prevent the collection of primary data.

One important observation from the interviews was that there seems to be a misperception among some in OAQPS about what secondary use of data entails. In some cases, for example, existing data that is generated by other groups within EPA may be perceived as primary data that has already been submitted to quality assurance, so that no further review is necessary EPA-generated data are perceived to be more credible than industry-generated data.

Sources and types of existing data. The sources of existing data for OAQPS include other groups within EPA or within OAQPS, reliable national sources, and industry. The existing data may be emissions data, ambient air data, air pollution control device performance data, census data, human activity data, meteorological data, housing surveys, and hospital admissions data.

Quality of secondary data. In general, the secondary users of existing data have assumed that the primary data generators have properly validated and verified the data before turning the data over to the secondary users. In general, OAQPS does not evaluate the quality of the existing data.

QA Project Plans. Some OAQPS staff may not be aware that the policy regarding secondary data has changed. That is, they may still perceive QA Project Plans as pertaining only to primary data collection and not to secondary uses of existing data, because that is consistent with the previous policy. They may perceive the development of QA Project Plans to be an administrative impediment. Due to the policy, the OAQPS staff will need to start preparing QA Project Plans for secondary uses of existing data.

Data evaluation procedures for existing data. In general, there are no formal procedures for evaluating the quality of existing data. Some informal data evaluation procedures are transmitted orally from one data evaluator to the next.

Metadata for existing data. In general, existing data used by OAQPS are not accompanied by metadata that characterize data quality.

Training and guidance. OAQPS conducted training for the preparation of QA Project Plans in accordance with EPA QA/R-5 in 2001 prior to the establishment of the OAQPS Secondary Data Policy. This training needs to be updated to address this policy. Many staff did not participate in the 2001

training because they did not see the course's relevance to their work.

Training and guidance on the new policy for the secondary use of existing data are seen as being desirable. This training could be presented as part of a project management or contract management course. The information may be more easily accepted by trainees as "data evaluation procedures" than as "quality assurance procedures." Any examples that are presented during training must be relevant to the divisional staff.

Authority. The divisional QA staff do not have sufficient authority to implement the secondary data policy alone. The commitment of division directors and senior managers to the policy is needed if it is to be implemented successfully.

Vulnerabilities related to OAQPS Secondary Data Policy

OMB guidelines. The new data-quality guidelines may result in greater level of external scrutiny of its procedures secondary uses of existing data by OAQPS for rule-making. The policy can be seen as a pro-active step in the direction of meeting the OMB IQ policy. However, the implementation of this policy will be a challenge.

Documentation. To be responsive to the OMB guidelines, OAQPS may need to develop adequate documentation of the quality of disseminated data and of the procedures that OAQPS or primary data collectors use to characterize data and evaluate data quality. OAQPS may need to document that it has considered data quality in its decision-making process.

QA Project Plans for Secondary Uses of Existing Data. The preparation of QA Project Plans for projects that involve secondary uses of existing data is a logical starting point to categorize projects according to the four-tiered approach, to implement the secondary data policy, and to document data quality. Unfortunately, the OAQPS staff may not be familiar with the preparation of QA Project Plans or with data evaluation procedures. They may be slow to prepare QA Project Plans and to evaluate data without management commitment and without effective guidance and training. The organizational culture may need to be modified to implement the policy.

Life cycle of data. Systematic planning must address the whole life cycle (i.e., generation, analysis, use, and dissemination) of existing data. When multiple groups within OAQPS or EPA are involved with different stages of the data's life cycle, each group may assume incorrectly that the other groups are evaluating data quality and may not take responsibility for evaluating data quality. Closer coordination between data generators, data users, and data disseminators is needed to ensure that data quality is characterized, evaluated, and documented before dissemination.

Scarce data. It may be difficult for OAQPS to implement its secondary data policy for projects where the scarcity of existing data is a larger concern than the quality of the existing data. That is, the data quality objective (DQO) process may not be an appropriate technique when the existing data or their metadata are scarce and there are not enough funds or expertise to collect primary data. Documenting the quality of scarce existing data may be more important than evaluating their quality.

Metadata. OAQPS structures for the dissemination of environmental data need to include metadata that characterize data quality. OAQPS staff need to report data quality throughout the life cycle of primary data and existing data from their collection to their dissemination.

Training and guidance. Project staff need effective training and guidance regarding procedures to prepare QA Project Plans for secondary use of existing data and to evaluate data quality. Project staff need guidance to evaluate the quality of existing data in the absence of adequate metadata. Furthermore, because individuals who deal only with secondary use of existing data may not be aware that training on QA Project Plans is now relevant to them, special effort must be made to make them aware of their need for this training.

Annual Recertification Program for Audit Standards used in the EPA $PM_{2.5}$ Performance Evaluation Program

Robert S. Wright and Jeffrey S. Nichol, RTI,
Michael L. Papp and Paul W. Groff, U.S. EPA
Michael W. Tufts, ARCADIS Geraughty & Miller

This paper describes procedures used to perform 152 annual recertifications of temperature, pressure, and flow rate audit standards. It discusses the metrology laboratories and the uncertainty of their recertifications. It describes the database for the standards that tracks their recertifications and shipments. Finally, it presents some illustrative recertification results and describes what these results reveal about the audit standards and the recertifications.

EPA $PM_{2.5}$ Performance Evaluation Program

EPA's Office of Air Quality Planning and Standards (OAQPS) established the Performance Evaluation Program (PEP) to evaluate the total measurement system bias of inhalable particulate matter ($PM_{2.5}$) samplers at State and Local Air Monitoring Stations (SLAMS) and National Air Monitoring Stations (NAMS).[1,2] The strategy is to collocate a portable PEP $PM_{2.5}$ sampler with a site's $PM_{2.5}$ sampler and to operate both samplers according to the Federal Reference Method. The exposed PEP filter is weighed at an EPA laboratory and the calculated PEP $PM_{2.5}$ concentration is compared to the corresponding value for the site's $PM_{2.5}$ sampler. The acceptance criterion for the agreement between the two concentrations is ±10 percent.

Environmental Services Assistance Team field scientists located in ten EPA Regional Offices visit SLAMS and NAMS sites at regular intervals for the collocated measurements. During the visits, the field scientists perform verifications of each $PM_{2.5}$ sampler's leak rate, ambient and filter temperature, barometric pressure, and flow rate using audit standards that are recertified annually.

Annual Recertification Program for Audit Standards

OAQPS requires annual recertifications for the audit standards to ensure their traceability to the National Institute of Standards and Technology (NIST). Acceptance criteria for the temperature, pressure, and flow rate standards are ± 0.5E C, ± 5 mm Hg, and ± 2 percent, respectively. Since Fiscal Year 2000 (FY00), recertifications have been conducted by four metrology laboratories, including one operated by EPA's Air Pollution Prevention and Control Division (APPCD).[3]

RTI handles the logistical and record-keeping aspects of the program[4,5]. Two rounds of recertifications are scheduled each year to allow the field scientists to have at least one of each type of audit standard at all times. RTI collects the standards from the field scientists and then ships them to the metrology laboratories. After the standards are recertified, they are shipped back to the field scientists.

The Audit Standards

Most audit standards are used by the field scientists and some are retained by OAQPS as spares. The audit standards and their manufacturers, model numbers, and numbers in the PEP inventory are given in Table 1. A total of 152 recertifications were performed in FY02. Digital thermometers/probes and flow transfer standards (FTSs)/electronic manometers were recertified as systems. Electronic manometers were also recertified separately. Primary flow meters were not recertified until FY02 because they were purchased in FY00 with two-year certifications.

Table 1. Manufacturer and Model Numbers of Audit Standards

Type of Instrument	Manufacturer	Model Number	Number
Digital thermometers	Control Company	4000	27
Temperature probes	Control Company	61220-604	27
Pressure verification devices	Druck	DPI 705	15
Pressure calibration devices	Meriam Instruments	LP200I	13
Digital pressure gauges	Psi-Tronix	PG2000	12
Primary flow meters	BIOS International	DC-40K	11
Flow transfer standards	Chinook Engineering	Streamline FTS	31
Electronic manometers	Dywer Instruments	Series 475 Mark III	30
Dry gas meters	Schlumberger (Thermo Andersen)	Galus 1.6	13

The Metrology Laboratories

The four metrology laboratories that have participated in the program are listed in Table 2. Each laboratory maintains NIST-traceable reference standards that are used to recertify the audit standards. The acceptance criteria for the audit standards and the uncertainties of the recertifications are given in Table 3. Each metrology laboratory recertified the audit standards at several points over the $PM_{2.5}$ sampler's normal operating range.

In FY00, the APPCD metrology laboratory recertified temperature and pressure audit standards, but it did not have a low-uncertainty reference standard for flow rate recertifications. The flow rate audit standards were recertified in FY00 and FY01 by commercial metrology laboratories. In FY02, the APPCD metrology laboratory acquired a low-uncertainty flow rate reference standard and all recertifications are now performed there. This arrangement reduces shipping and recertification costs and allows a uniform format for certificates. It also reduces the uncertainty of flow rate recertifications.

The Recertification Database

RTI developed a Microsoft Access® relational database to document audit standard recertifications and to track the standards as they are being recertified. The database maintains an inventory of the audit standards, model and serial numbers, and recertification and operational statuses. It records information about field scientists, shipping addresses, shipment tracking numbers, metrology laboratories, and reference standards that were used for the recertifications. The database records comments by the field scientists and the metrology laboratory about each audit standard. It generates inventory reports for each EPA Regional Office, recertification

Table 2. Metrology Laboratories that Recertified $PM_{2.5}$ PEP Audit Standards

Metrology Laboratory	FY00	FY01	FY02	Audit Standard Recertified
APPCD Metrology Laboratory U.S. Environmental Protection Agency Mail Drop 91 86 T. W. Alexander Drive Research Triangle Park, NC 27711	U	U	U	Digital pressure gauges Digital thermometer/probes Dry gas meters Flow transfer standards Electronic manometers Pressure calibration devices Pressure verification devices Primary flow meters
Colorado Engineering Experimental Station (CEESI) 54043 County Road 37 Nunn, CO 80648		U		Flow transfer standards Electronic manometers
Chinook Engineering 555 Abasaraka Sheridan, WY 82801	U			Flow transfer standards Electronic manometers
Thermo Andersen 500 Technology Court Smyrna, GA 30082-5211	U	U		Dry gas meters

Table 3. Comparison of Acceptance Criteria and the Uncertainty of Recertification

Audit Standard	Acceptance Criterion	Reference Standard Type	Uncertainty of Recertification
Temperature	±0.5EC	Platinum resistance thermometer	±0.2EC (Model 4000) (APPCD) ±0.1EC (Model 61220-604) (APPCD)
Pressure	±0.7%[a]	Dead weight piston	±0.10% full scale (APPCD)
Flow Rate	±2%	Spirometer (bell prover) Critical flow venturis Critical flow venturis Laminar flow element	±1.5% (Andersen) ±2% (Chinook) ±0.5% (CEESI) ±0.5% (APPCD)

[a] The acceptance criterion (±5 mm Hg) given in the PEP QAPP is ±0.7% at 1 atmosphere (760 mm Hg).

reports for each audit standard, and packing slips for each shipment.

Illustrative Examples of Recertification Results

Figure 1 displays the recertification results for 15 Druck pressure verification devices recertified in FY01. The acceptance criterion for the devices is ±5 mm Hg. The as-received errors (before recertification) for 13 of the 15 gauges met this criterion. Some devices' readings were in error by as much as ±8 mm Hg. The metrology laboratory adjusted the devices' calibrations, if necessary. The as-left errors (after recertification) for all devices met the acceptance criterion. This example demonstrates the necessity of periodic recertifications of audit standards to maintain their accuracies. The receipt of 2 devices that did not meet the criterion suggests that more frequent recertifications and/or field verifications of these devices are needed.

The FTS is an orifice and the pressure drop across it is measured by an electronic manometer. The calibration equation is given as the slope and intercept of a line relating flow rate to the square root of the pressure drop. To estimate year-to-year calibration shifts, the $PM_{2.5}$ sampler's design flow rate (16.7 L/min) is converted to the equivalent pressure drop using one year's calibration equation and then that pressure drop is converted to the corresponding flow rate using the next year's calibration equation. The design and shifted flow rates can be compared to estimate the calibration shift.

Three sets of certification data are available for 23 FTSs and an incomplete set is available for 12 FTSs in FY02. As shown in Figure 2, most FTSs met the ±2 percent acceptance criterion over the 3-year period. Four FTSs did not meet it at some point. Comparison of the yearly calibration shifts for two FTSs suggests that a problem occurred during their recertifications in FY00. The calibration of one shifted by -4.4 percent from FY99 to FY00 and by +4.1 percent from FY00 to FY01. Another FTS's calibration shifted by -3.3 percent and by +4.2 percent, respectively.

No errors were evident from inspection of the two FTSs' recertification certificates, but analysis of all FTS recertification data revealed that the pressure differential readings for these two FTSs were noticeably greater than the corresponding readings for other FTSs in FY00, but not so different in FY99 or FY01. The problem may not have been detected by the metrology laboratory because the recertification certificate reports the slope and intercept, which may be hard to interpret. This example illustrates the advantage of analyzing recertification data for multiple audit standards across multiple years to detect long-term trends in the calibrations of the standards that would be difficult to detect if only data from individual recertifications are reviewed. It also illustrates that one cannot always assume that any recertification by a metrology laboratory has been done correctly.

Field and Laboratory Checks of Recertified FTS Audit Standards

RTI conducted a field check of 27 recertified FTSs in FY01 to double-check the recertifications. They were compared with the internal flow meter of a $PM_{2.5}$ sampler and a BIOS primary flow meter at a SLAMS site. The agreement between the FTS and the internal flow meter's readings ranged from -0.3 to 1.6 percent with a mean of 0.6 percent. The agreement between the FTS and the BIOS primary flow meter's readings ranged from -2.0 to 0.7 percent with a mean of -0.8 percent. These agreements support the belief that the FTSs' recertifications attained the ± 2 percent acceptance criterion in FY01. A laboratory check in FY02 of 12 recertified FTSs against a flow rate reference standard yielded a mean agreement of -0.4 percent with values ranging between -1.0 and 0.4 percent.

Advantages of the Annual Recertification Program

The main advantage of the annual recertification program is that it enhances the defensibility of PEP audits and $PM_{2.5}$ measurements by eliminating one source of measurement uncertainty. The centralized recertification of all audit standards helps to ensure that $PM_{2.5}$ measurements across the country are comparable and traceable to NIST. Additionally, there is an economy of scale advantage associated with recertifying many audit standards in the same metrology laboratory.

The database provides documentation that audit standards remain traceable to NIST. It allows EPA to monitor their certification status and helps to ensure that they are recertified at proper intervals. It tracks shipments to ensure that each standard is returned to the correct EPA

Regional Office.

The examples illustrate that periodic recertifications of audit standards is needed and that the program helps to identify standards that do not meet acceptance criteria or that are prone to error. The analysis of recertification results for many audit standards across several years allows EPA to track long-term trends in the audit standards' calibrations. It helps EPA to verify that recertifications are done correctly.

References

1. U.S. EPA, *Quality Assurance Guidance Document 2.12, Monitoring $PM_{2.5}$ in Ambient Air Using Designated Reference or Class I Equivalent Methods*, November 1998 (available at http://www.epa.gov/ttn/amtic/files/ambient/pm25/qa/m212covd.pdf).
2. U.S. EPA, *Quality Assurance Project Plan for the $PM_{2.5}$ Performance Evaluation Program*, February 1999 (available at http://www.epa.gov/ttn/amtic/files/ambient/pm25/pmqa.html).
3. Groff, P.W., *A Solution for the Need to have Defensible, Documented, Quality Data Traceable to an Accredited Laboratory at Low Cost*, presented at the 20th Annual National Conference on Managing Environmental Quality Systems, 2001.
4. Nichol, J. S., *Recertification of Calibration Standards for the EPA $PM_{2.5}$ Performance Evaluation Program for Fiscal Year 2000,* RTI/7503.OP1.201/-01F, May 2000.
5. Nichol, J. S., *Recertification of Calibration Standards for the EPA $PM_{2.5}$ Performance Evaluation Program for Fiscal Year 2001*, RTI/7503.002.303/-01F, May 2001.

Disclaimer

The information in this document has been wholly or in part funded by the U.S. Environmental Protection Agency under Contract Number 68-D-98-032 to RTI. It has been subjected to the Agency's peer and administrative review and has been approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Figure 1. FY01 Recertification Results for Druck Devices

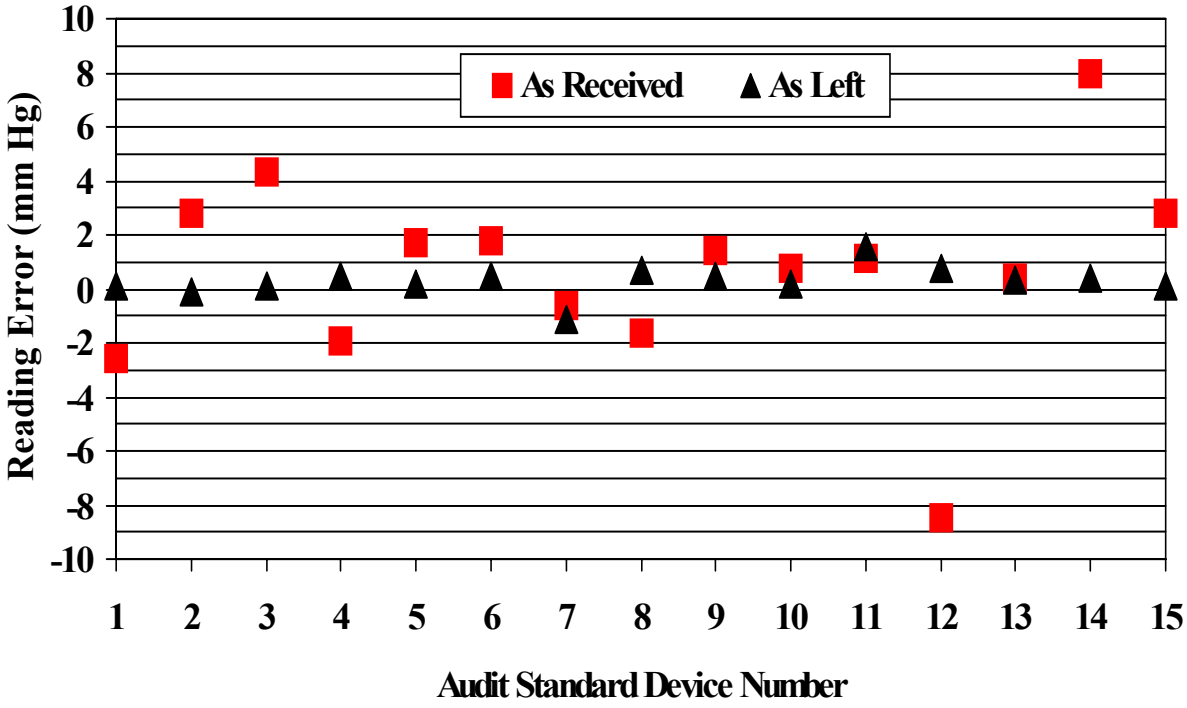
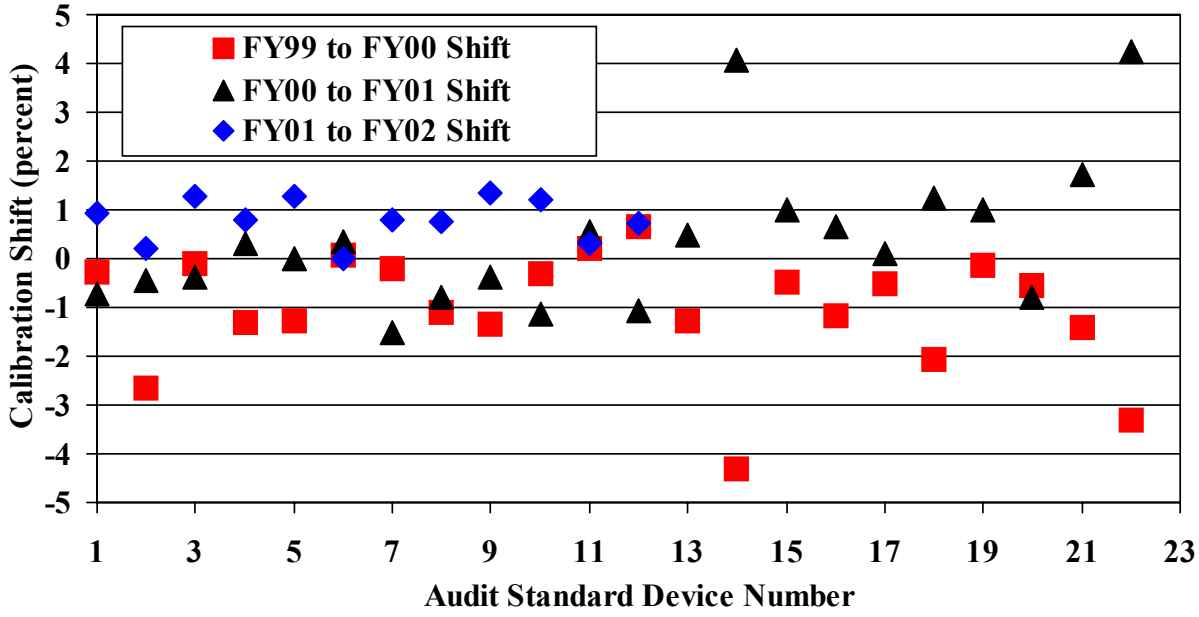


Figure 2. Year-to-Year FTS Calibration Shifts



Data and Metadata Reporting Standards for the U.S. Environmental Protection Agency's PM Supersites Research Program[1]

Les A. Hook, NARSTO Quality Systems Science Center, Oak Ridge National Laboratory[2]
Sigurd W. Christensen, NARSTO Quality Systems Science Center, Oak Ridge National Laboratory
William B. Sukloff, Environment Canada, Meteorological Service of Canada

The EPA Supersites Research Program needs consistency of metadata and data structures to facilitate information sharing among investigators, analysts, and ultimately secondary data users. Under the auspices of NARSTO[3] a successful mechanism was created to develop and implement reporting standards. The development effort included working closely with Supersites data coordinators, investigators, and technical experts, and also leveraging from existing data standards and practices. Overall, the standards are getting good acceptance from the atmospheric research community.

The U.S. Environmental Protection Agency is sponsoring a major atmospheric particulate matter (PM) data collection effort in seven major U.S. cities, called the PM Supersites Research Program (Fig. 1). The Supersites Program's objectives are to (1) characterize PM and its constituents, (2) collect data and samples to support health and exposure studies to reduce uncertainty in setting National Ambient Air Quality Standards, and (3) compare emerging sampling and analysis methods with routine techniques to enable a smooth transition to advanced methods. In addition to analyzing individual site PM and atmospheric conditions, the data from all the Supersites are to be capable of being integrated for cross-site analyses, and are to be archived in a timely manner and be readily available to the public.

Figure 1. U.S. EPA Supersites Research Program

Data reporting was addressed in the Cooperative Agreements that implement the Program. Data Coordinators support the data reporting process at each Supersite (Fig. 2). The NARSTO Permanent Data Archive (PDA) at the Langley NASA DAAC was designated the final repository. The PDA has a required self-documenting data format, the NARSTO Data Exchange Standard (DES), with several metadata requirements. The archiving process is mediated by the NARSTO Quality Systems Science Center (QSSC), Oak Ridge National Laboratory. NARSTO encourages scientists to document their data at a level sufficient to satisfy the well-known "20-year test". That is, someone 20 years from now, not familiar with the data or how they were

[1] Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency. Michael Jones, Project Officer.

[2] ORNL research was sponsored by U.S Environmental Protection Agency and U.S. Department of Energy and performed at Oak Ridge National Laboratory (ORNL). ORNL is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725.

[3] NARSTO is a tri-national, public-private partnership for dealing with multiple features of tropospheric pollution, including ozone and suspended particulate matter.

obtained, should be able to find data of interest and then fully understand and use the data solely with the aid of the documentation archived with the data[4]. Integration of data in future analyses demands consistently defined metadata elements and values.

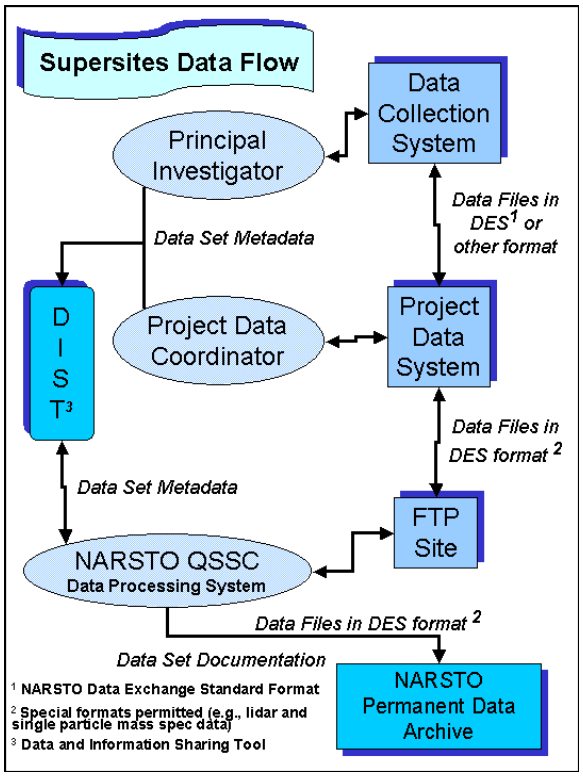
Data and Metadata Standards Development

We began the development of the data reporting standards began by working with the Data Management Coordinators for each Supersite. A Data Management Working Group (DMWG) was formed with the QSSC as the lead. The Working Group communicated through weekly teleconferences to deal with consistency of metadata content and data reporting format. Minutes of the teleconferences discussion and decisions were distributed to the DMWG and Site Principal Investigators. The continued support of the EPA Program Mangers is critical to the success of this effort.

We incorporated metadata elements and values from other metadata standards when available to promote consistency within EPA and the atmospheric research community, and to anticipate integrating data from additional sources. For example, in addition to existing NARSTO standards, we used site descriptors and event flags from EPA AIRS, detailed flags from EPA Region 5, use of the CAS Registry Number and CAS Index Name for chemical identification from EPA CRS, and the non-chemical variable naming syntax from the DOE ARM Program.

By leveraging existing resources and the developing data management resources and technical expertise of the individual Supersites and NARSTO, we were able to develop a set of robust reference materials and supporting systems. Site-specific implementation flexibility is always a consideration and was maintained when possible. Each metadata standard was completed within the DMWG and then sent to the Site Principal Investigators for approval, after which they were considered Consensus Metadata Standards. The DMWG updates these as needed.

Figure 2. Supersites Data Flow Diagram.



Consensus Metadata Standards

Site Identification: Identifies a standard syntax for naming fixed and mobile sites used by studies or networks for air quality sampling and monitoring. A site is assigned a 12-character site identifier that includes a four-character site abbreviation (the "site mnemonic"). The first four characters identify a study or network. A master list of site names from the Supersites program will be assembled by the QSSC. The master list would include additional information about each site, as available: latitude,

[4] National Research Council, Committee on Geophysical Data, *Solving the Global Change Puzzle, A U.S. Strategy for Managing Data and Information,* National Academy Press, Washington, D.C., 1991.

longitude, elevation, EPA AIRS identifier, land use, and location type.

Identifying Chemical and Physical Variables and Descriptive Metadata:

- Identifying Chemical Substances with a CAS Registry Number: Valid values are the CAS Number (with "C" prefix -- "C" prevents spreadsheet programs from converting some CAS numbers to dates) and Chemical Name (either CAS-9CI, IUPAC, or other common name). CAS numbers and preferred names for common atmospheric constituents are in a reference table.
- Identifying Chemical Substances, Calculated Quantities, and Physical/Non-chemical Measurements that do not have a designated CAS Registry Number: Variable names are formed beginning with the root concept, and followed by a detailed modifier if needed, separated by a ":". For example, PM10: area, PM10: count, PM10: mass, and Temperature: air, Temperature: dew point, Temperature: virtual. These variables can be method specific and require special differentiation. Definition of new variables is relatively straightforward when the format is followed.
- Identifying Metadata Elements: Valid variable names for metadata elements including site information, locations, dates, times, and sampling conditions are provided in a reference table. The correct format is the root concept, followed by a detailed modifier if needed, separated by a ":". For example, Date start: local time, Date end: local time, and Latitude: decimal degrees, Longitude: decimal degrees.

Data Quality Flags: Reported data values must be assigned at least one data quality flag by the data originator to indicate whether the data are valid without qualification, valid but qualified/suspect, or invalid due to serious sampling or analysis problems. These flags may be the NARSTO data qualification flags or other more detailed flags as defined by a Project. Project-defined flags must be mapped to NARSTO flags. Reference tables of NARSTO standard flags, detailed project flags, and EPA AIRS exceptional-event flags are provided for users.

Changes and additions to the reference tables are controlled. Site investigators and data users are encouraged to work with their Data Management Coordinators to suggest improvements in and additions to the reference tables. A Data Coordinator should recommend additions or changes to the DMWG and QSSC for discussion and consensus. Subject matter experts are consulted when appropriate.

Data Exchange Standard Development

Data files submitted for archiving should be in the NARSTO Data Exchange Standard (DES) format. The DES format follows a spreadsheet-compatible layout and is stored as ASCII comma-separated value (.csv) files. The DES does not rely on row position to identify metadata information, but uses tags to describe the information contained in the row. The DES is a self-documenting format with three sections: the header section contains information about the contents of the file and the data originator; the middle section contains metadata tables that describe/define sites, flags, and other codified fields; and the final section is the main data table that contains key sampling and analysis information and the data values.

The consensus metadata standards for site names, data quality flags, and parameter names, plus key characteristics (see below) are implemented in the DES. An Excel/97® template for inputting data and metadata has been developed to support data providers. The template is annotated with comments, instructions, frequently asked questions, and examples of completed

files. Within the template are picklists for selecting values for various metadata fields to promote consistency in terminology.

Supersite Enhancements to the Data Exchange Standard: Until the initiation of the PM Supersites Program, the DES had been used primarily with gaseous atmospheric constituents and meteorological measurements (e.g., ozone, air temperature, and solar radiation). The sampling and measurement of these constituents is generally straightforward with well-defined methods and reporting conventions. It soon became clear that PM measurements are not so easily characterized. PM results (e.g., size-differentiated mass, number, and chemical composition) need more metadata than just the name, units, and analysis method. In many cases, results are operationally defined by the specific field sampling configurations, measurement devices, and conditions, and the laboratory sample preparation and analysis methods.

To address the expansion of measurement types, a set of key characteristics (Table 1) was defined to capture enough of the measurement information to be meaningful and helpful in a data file, while avoiding excessive detail. The key characteristics are metadata fields that hold general descriptions of the field, instrument, and laboratory conditions. Detailed information would always be included as companion files, such as the Quality Assurance Project Plans. Key characteristics, metadata values, and organization of the DES were defined through invaluable interactions of Data Coordinators, with PIs and with other field and laboratory technical experts.

Key Characteristics provide general sampling and analysis information that describes the data.	
> OBSERVATION TYPE > SAMPLING HEIGHT (M AGL) > FIELD SAMPLING OR MEASUREMENT PRINCIPLE > INLET TYPE > MEDIUM > COATING OR ABSORBING SOLUTION/MEDIA > SAMPLING HUMIDITY OR TEMPERATURE CONTROL > PARTICLE DIAMETER--LOWER BOUND (UM) > PARTICLE DIAMETER--UPPER BOUND (UM) > PARTICLE DIAMETER--MEDIAN (UM) > WAVELENGTH (NM) > WAVELENGTH--LOWER BOUND (NM) > WAVELENGTH--UPPER BOUND (NM) > SAMPLE PREPARATION > LABORATORY ANALYTICAL METHOD	> VOLUME STANDARDIZATION > BLANK CORRECTION > INSTRUMENT NAME AND MODEL NUMBER > MEASUREMENT PRINCIPAL INVESTIGATOR > EXPLANATION OF ZERO OR NEGATIVE VALUES > EXPLANATION OF REPORTED DETECTION LIMIT VALUES > DETECTION LIMIT VALUES > EXPLANATION OF REPORTED UNCERTAINTY VALUES > UNCERTAINTY VALUES **Picklists for selecting Key Characteristic values are included in the DES template.**

Table 1. Key Characteristics Included in the Data Exchange Standard.

Additional Data Reporting Guidance

To ensure that data can be integrated for successive analyses, consistently reported data and metadata are essential. Supersites' Technical and Quality Assurance Leads provided this guidance.

Submittal of Uncertainty Estimates: EPA is strongly recommending that within each Supersite the research investigators and data managers estimate and report the data uncertainties.

Estimating the uncertainty of the data collected is of paramount importance to the purpose of the Supersites Project. Data users will need to understand the uncertainty of the data, which will enhance confidence in their assumptions and predictions. The DES has been updated to ensure that uncertainty can be conveniently reported by data providers and can be interpreted by data users.

Data Reporting Conventions Guidance: To further promote consistency among data products, the "level" of data to report has been specified for data providers. Mass/volume measurements (e.g., from filters) should be reported as concentrations, rather than separately as mass and volume. PM mass data should be referenced to local ambient temperature and pressure conditions to be comparable to federal reference method PM data. Associated meteorological data, including temperature and pressure conditions, should be reported either in the same file or in a referenced meteorological data file.

Similarly, units for chemical variables and particle measurements are specified to follow SI standards, when possible, or units commonly used by the research community. A pick list of units has been implemented in the DES.

Data Archiving Process Guidance

The Supersites Data Managers are provided with specific guidance for carrying out the final steps in the data management and archiving process. Specifications for data set and data file naming and configuration control are given. The NARSTO Data and Information Sharing Tool has a convenient metadata entry/export feature for efficient preparation of archive documentation. The QSSC is the source for information and assistance, and it verifies submitted DES data file format compliance and mediates these archiving activities.

Read and Verify Program Verifies Data File DES Format Compliance: Data Coordinators send the completed DES files to the QSSC. A Read and Verify Program checks numerous format and content elements of the DES files by verifying that key characteristic values are in the picklist reference tables, key phrases are correctly formed, variable formats and format types are correct, CAS numbers are in reference table, dates and times are properly formatted, the UTC time offset is correct, flags are in the flag look-up table, and sites appearing in the main data table have corresponding entries in the site information table. The Read and Verify Program also calculates and inserts into the file a set of summary statistics records for each numeric variable in the file. The statistics include minimum, maximum, mean, standard deviation, n, number of missing values, and total number of records. As an additional QA check, time series plots can be created. The plots include summary statistics and key characteristic values. (Fig. 3).

If the Read and Verify Program finds problems, a QA problem report is sent back to the data originator, with some guidance on how to correct the problems. The data originator should correct the problems and resubmit the data file. This process continues until all parties are satisfied with the dataset. Supersite Data Coordinators with SAS® software are also running this program before submitting files to the QSSC.

Data and Information Sharing Tool (DIST) Generates Archive Documentation: The NARSTO DIST was implemented for Supersites to support compiling data set metadata and generating archive documentation. Either a Data Provider or the QSSC can enter metadata into the DIST

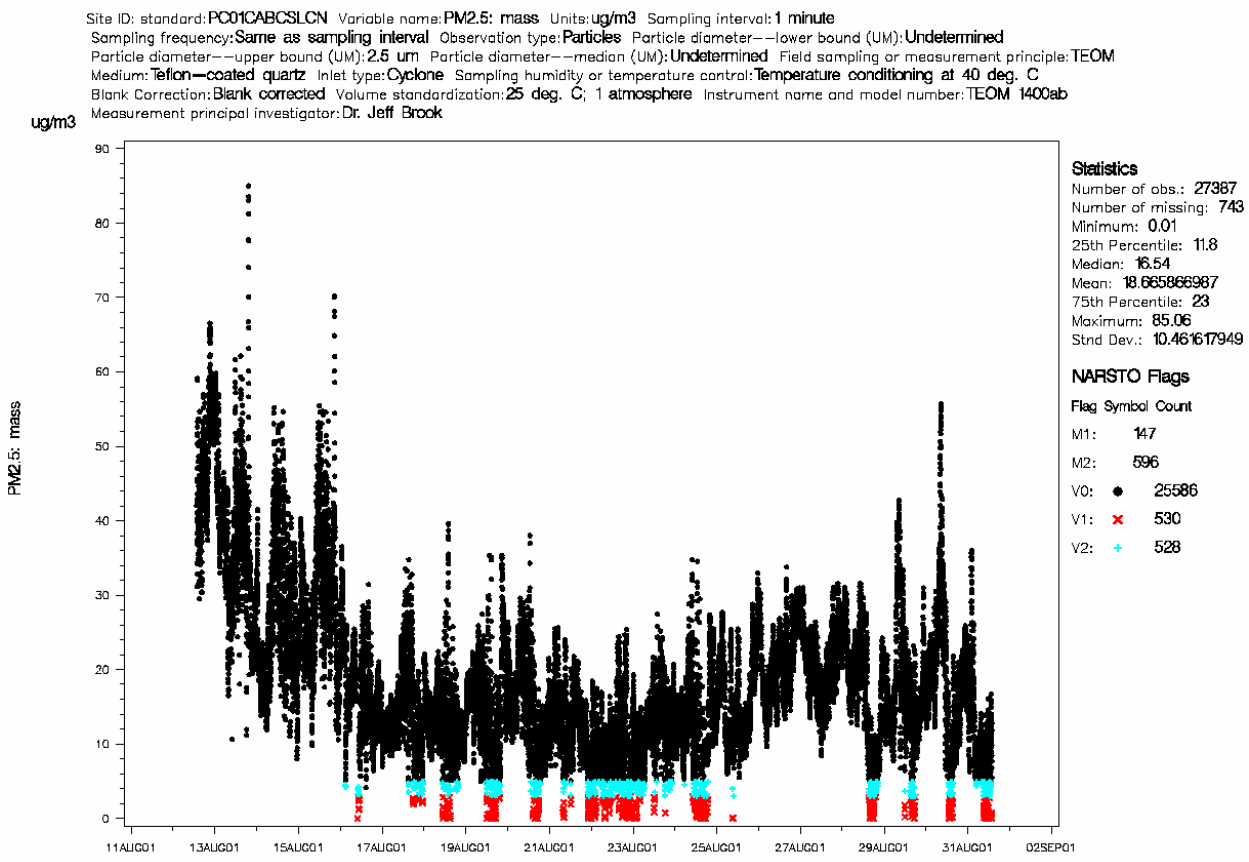


Figure 3. Example Time Series Plot Generated by Read and Verify Quality Assurance Program.

metadata editor and output it in the formats needed by the NARSTO Permanent Data Archive. When documentation is complete and the QSSC data file verification checks are complete, the properly formatted data files can be moved to the archive.

All of the standards, the DES template, DIST, and guidance documents referenced in this paper can be accessed through the QSSC web site [http://cdiac.esd.ornl.gov/programs/NARSTO/].

Conclusions

The accomplishments of this development effort to date are, in no small part, due to the early recognition of the need for data management planning and implementation and its inclusion in the Cooperative Agreements, and the continued support of the EPA Program Mangers. This standards development process was successful in integrating the input of Supersites data management and research staff with existing NARSTO and other applicable standards. The product is a robust set of Supersites data and metadata reporting standards that will facilitate current PM data reporting, analyses, and archiving activities: can be extended to additional data types; and will support integrated analyses and future research projects.

Building the Airplane in Flight:
An Auditing Approach to Quality Management System Development

Malcolm C. Burson, Maine Department of Environmental Protection

In June of 2000, Maine DEP (in company with the other five New England states) found itself under EPA pressure to develop and document a quality management system by the end of the calendar year. In the frenzy that followed, the Department decided to use a private sector model for system development that called for a prospectively-focused QM plan that would be implemented through formal auditing. That is, instead of trying in advance to identify all the deficits in departmental quality management, and then assigning individuals and task groups to create structures to fill the gaps, Maine's QMP specifies the future desired system in broad terms. ME-DEP then uses its own cadre of trained auditors to assess current practice against the "condition expected" in the QMP, knowing that in many cases basic quality management practices will need to be developed. This approach assists program managers, particularly (but not exclusively) in areas sensitive to environmental data standards, in creating systems and practices that are rooted in reality, and that are perceived to add value to the Department's core work instead of just additional paperwork. Audit reports create a continuous feedback loop assuring that written procedures document actual operations. Finally, the results of auditing identify areas in which the QM system (and its plan) should be elaborated or refined, leading to an iterative process by which quality approaches are infused in all areas of DEP operations. In the twelve months following EPA-Region I's initial approval of Maine's QMP, a total of seven audits were completed at various program levels, including two focused on critical QMP elements: Documents and Records; and Computer Hardware/Software.

This presentation will:

- *Briefly describe the background and implementation of this approach;*
- *Identify some of the factors which led to success;*
- *Describe, using selected examples, some of the early outcomes of the program.*
- *Time will be allowed for questions.*

The title of this presentation will perhaps be familiar to you from a nationally-telecast commercial advertising the consulting services of the same company that brought us "herding cats" and "running with squirrels." Visually, it presents a large fast-moving airliner, with teams of designers and fabricators tuning the engines and adding aluminum skin at 30,000 feet, all the while looking into the aircraft at the people going about their business.

That's very much the story of how Maine DEP continues to develop and implement its quality management system. And as much as I love telling the story of how this came to be, our time is limited, so I'll be providing you only with highlights. Further, since I'm not sure how familiar you are with the process of quality auditing, I'm going to allow lots of time at the end for questions that may delve into the

mechanics of our approach. So let me begin by providing you with the headline that started all this (slide 1). The letter announcing this news was dated June 12, 2000, and the deadline for a draft Quality Management Plan was December 31. We didn't panic…quite. But we quickly faced three critical decisions (slide 2). We could, as some states have done, choose to limit application of the quality management system only to Federally-funded environmental data operations. We chose instead to start with a system that would encompass all areas of operations, principally because we thought we would waste more energy figuring out what parts of work to carve out then we would gain from using less-rigorous standards in part of the Department. And frankly, we believed that was just the wrong thing to do in any case.

Decision 2 was whether to figure out how to design and document a Quality Management System by ourselves, or find assistance. We chose the latter, even though it meant finding funds (some through re-negotiation of the PPA) we would rather have spent on "real" environmental work. As it turned out, this choice paid for itself many times over in not having to take one or more talented staff away from their primary obligations in order to accumulate documentation, design a system, and draft a plan to meet the deadline. And our consultants, who had private sector ISO 9000 and 14000, and EMS development experience, were jewels.

But it was decision 3 that made all the difference. One of our senior managers, then recently arrived from the private sector, was himself a certified ISO registrar with significant experience in QMS development. He pointed out that in most organizations, the run-up to QMS implementation or ISO certification involves an organization-wide frenzy of gap analysis, followed by the creation of endless lists of SOPs to be completed, documentation to be accumulated, flowcharts to be created, training to be required, and then collective hand-wringing over the extent to which such a paper chase will drive people crazy. Instead, he suggested, we could "work from the other end" as it were, by creating a QM Plan that outlined the future desired state of the system as a whole. An auditing methodology could then compare the *existing* system in place with the desired future, using the ANSI standards as a benchmark. The results of a given audit would not only provide guidance to operating programs on areas needing improvement, but would also feed back to revise and improve the QM system itself. Most importantly, audits would focus on assuring and improving quality *as work is actually done*, rather than in abstract, and would thus add value at the program level. In other words, auditing is seen not primarily as a means to look over people's shoulders and tell them what they're doing wrong; instead, it's a way to build learning for everyone.

Fortunately, we were blessed with contacts in the EPA-New England quality office who took a very open-minded attitude to our work. They hadn't heard of a QMS and Plan being created in this manner, but they were incredibly supportive along the way. And as a result, we became the first New England state to have an approved QMP, and have subsequently become the first to receive delegated authority for QAPP approval. (6)

So with the basic structure of the airframe in place, we took off. Our consultants provided four days of lead auditor training to a cadre of 16 staff from throughout the Department. The experienced senior manager I spoke of above pointed out that developing internal auditors had multiple benefits. Not only would we be able to carry out a number of second-party audits that would build the system, but the knowledge auditors gained about assuring and managing quality through inquiry outside their particular program area would be <u>immediately</u> transferable to their own units. Thus, when the audit regime worked its way around to that unit, they would tend to be far better prepared.

Maine DEP's QM system is overseen by a Steering Committee of six senior and middle managers, assisted by a coordinator in each major bureau. We believe that our shared leadership approach to top-level quality management has not only been a good use of resources, but has allowed our progress to be more visible in the Department than had we carved out a stand-alone office of quality. Among the

Committee's responsibilities is to set priorities for auditing, based on several principles. First, we try to identify areas of risk or vulnerability, where environmental data quality in particular must be as defensible as we can make it. But we balance this by asking program areas if they believe being audited will add value to their work, and by making sure that we spread the pain to some extent. (Indeed, the manager of the program area that found itself included in parts of two different audits in slightly over 2 months received our "Quality Camper of the Year" award in recognition of his unit's sacrifices!)

We also quickly identified three different levels of auditing (slide 3). At the highest altitude, we know that we need continuously to audit the QM system itself, using the QMP as the metric. So far, we have done this by auditing a particular QMP element across the Department, using a sampling approach to make sure we're learning enough about, for example, Documents and Records processes to be able to generalize the improvement response. At middle levels, we've found program areas more than willing to participate in what we call "process audits." In this case, the scope is determined by the boundaries of a particular work process, and the goal is to look at all aspects of quality management within that process. In our first audit of this sort, the team reviewed the flow of work associated with the EPA Permit Compliance System, involving as this did the licensing, compliance, and data management functions that in our Department across a number of program boundaries. Finally, down at a few thousand feet, where one can focus on specifics, the program carries out compliance audits on existing discrete QM systems, generally things like existing QAPPs. By maintaining a balance among these three levels of inquiry, we try (as has been a watchword for us since the beginning) to keep the intrusiveness of the auditing regime below the radar screen for most people most of the time. And this strategy also allows us to bring a "just in time" approach to building QM System awareness in the Department. Most people have had their first involvement with the system in the run-up to an audit. They then are immediately involved in the learning and quality improvement activities that result, instead of sitting through a half-day "welcome to QMP" workshop that has little value until it can be acted on in the course of their environmental work. (10)

I turn now to tell you of some of our experience to date, and some of the outcomes. We surveyed all affected managers after completing our first four audits, to find out what their experience had been. Perhaps their most interesting response was the report that for the most part, they were not surprised by the corrective action requests or auditor observations that resulted. That is, auditing tends to confirm and validate what managers already know about aspects of their operations that need improvement in order to meet the expectations of quality. This approach has the advantage of providing a concrete, actionable link between a current way of operating, and the desired or required state. This goes hand-in-hand with our approach of soliciting audit "volunteers." Since most managers realize that the QM system is going to reach out and touch everybody eventually, and most also understand the increasing importance of assuring environmental data quality, our approach provides them the opportunity to engage more formally what they already knew they would have to do.

We have learned that for auditing to be successful, excellent advance work is a necessity. But this also supports the eventual goals. The mutual process by which an audit team and a program group prepare for, and carry out, an audit creates learning for all concerned. Auditors from a totally different program area become intimately familiar with how a program or unit works, and they report that they carry examples of "best practice" back to their home units. At the two higher levels, audit preparation by the program almost always requires thought about just how processes work, whether they're currently documented or not. In addition, the gathering of documentation for advance review immediately uncovers precisely the sorts of improvement opportunities that managers have always known about, but which previously hadn't had much immediacy. Finally, the mutual process of determining the scope of a given audit demands that both auditors and auditees think carefully about issues of quality throughout the program. (12)

The introduction or development of a quality management system is probably never greeted with loud

hurrahs and infectious enthusiasm on the part of most staff. After all, they quickly realize that they will either have to do things differently than they have before; or they will have to add something to already full plates, or (most likely) both. But we quickly discovered some natural allies, whom I would encourage those of you charged with this responsibility to enlist from within your own organization. These are the folks who have been operating on the basis of documented QA/QC requirements all along. In our case, this was the gang of highly-independent, management-averse, not-to-be-trifled-with cowboys and cowgirls in our Superfund and uncontrolled sites unit. As it happens, we carried out our first compliance audit of their QAPP, and discovered people who for all their cynicism had developed a document that they were proud of, and more importantly, lived by on a day-to-day basis. They "get" QA/QC, and quality management, and will grudgingly admit, when pressed, that while planning for and implementing quality in their unit was a pain in the posterior, they know it works for them. The audit produced the expected relatively minor requests for corrective action (it was as much an exercise in on-the-job training for auditors as anything else), with corrective action plans carried out and closed expeditiously.

It's the follow-up to these CARs and CAPs that produces real challenges for us. Since we are committed to directing audit results to the managers and staff closest to the operational level involved, and rely on a tracking system and bureau coordinators (each of whom has at least 14 other higher-priority responsibilities) to oversee implementation, we haven't yet been entirely successful at riveting audit response fully onto the frame of the organization. We expected this: after all, presenting an already-busy manager with a list of 10 new projects usually results in a "deer in the headlights" look. We give each recipient of a CAR four weeks to develop a plan for corrective action, and, if the "fix" involves complex or systemic change, we know it's going to take a while to carry out. We encourage managers to include a project time line or bench marks as part of their plan, but we also know that within a short period of time, whatever becomes the next important thing is likely to drive an internal correction to a system well down the priority list.

As I noted above, we try to be sensitive to the issue of overload in choosing where audits will take place, and the same is true for CARs. Lead auditors are instructed to limit their requests to ten or fewer carefully chosen action areas, and to try to balance those likely to need large-scale intervention with "low hanging fruit." After all, we know that finishing this airplane will take more than one round of audits and resultant actions. And beyond the program level, we've had to become particularly careful with how we respond to bridging the gap between "condition expected" and "condition found" when we audit at the department-wide QMP level. For example, our initial documents and records audit produced the clear and urgent need for developing and implementing department-wide documentation and record-keep standards, and thus an infrastructure to support it, and eventual change in everyone's work processes. This will take several years to make even partially real, and we're taking our time before we institute an audit that will likely have the same effect in some other area of operations. We're already hearing significant grumbling at the most senior level that "this QMP thing" is disrupting what's perceived to be important day-to-day work.

For those of you who may be interested in building your own system in this way, we would also say that maintaining the internal auditor cadre can be a challenge. We've had several trained auditors who, through promotion, are bringing their commitment to quality management into a new place in the organization, yes, but who are unfortunately no longer able to commit the time necessary to being on an audit team. We're also wrestling with how we assure that we continue to recruit people who will bring the right attitudes to the special task of being an objective auditor. We have been realistically clear to potential auditors (and their supervisors) about the time expectations involved, but of course when one's primary tasks are filling the plate, it's not easy for an auditor to commit to a newly-formed team that will significantly affect his work for the next 4-8 weeks.

Yet even with these and other challenges, we continue to believe that we made the right decision. We've

already accomplished far more in the first year of our QM system than any of us believed possible. As participating in audits touches more and more of our 425 staff, we're seeing people beginning to anticipate: that is, not waiting for the audit to find them before putting effort and resources into things like documenting fundamental processes, or taking on the long-delayed task of assuring operational consistency across regional boundaries.

Finally, what's discovered in our audits not only provides Corrective Action Requests and Plans that mean improving quality in the area audited; it will also continue to provide information that will be needed to modify the Department's QM Plan so that it better accords with reality. For in our experience, there will always be well-intentioned requirements in the QMP that are discovered to be thoroughly impractical at the program level. We understand the QMP as a living structure which itself is continually modified in flight in order to reach the distant quality airport as efficiently and effectively as possible.

Quality Performance Evaluation:
Measuring Quality Management Systems Performance

Gary L. Johnson, U.S. Environmental Protection Agency

BACKGROUND

Organizations in private industry and in government world wide have implemented and used quality management systems for more than 50 years to produce products and services that meet their customer's needs and expectations. Many of these quality systems are based on international consensus standards, like ISO 9001:2000, or national consensus standards, like ANSI/ASQC E4:1994 for environmental programs, which provide consistent, standardized requirements for planning, implementing, and evaluating quality control (QC) and quality assurance (QA) practices in their operations and activities. Audits enable users to determine conformance of their quality systems to the requirements of the standards applied. Since management systems do not typically have performance specifications like measurement devices or laboratory methods, the ability of managers to assess the performance of their quality management systems has been limited. However, performance of the management system is a key factor to an organization's overall mission performance and should be included in management review processes.

INTRODUCTION

Quality Performance Evaluation (QPE) is a new tool which can help managers assess the performance and effectiveness of their quality management systems. QPE is based on setting goals and objectives for the quality system, identifying specific performance measures for those objectives, and using performance indicators to determine if the objectives are being met. Typically, the performance objectives are specific to an organization and its mission. For example, a quality system for a commercial analytical laboratory will have different objectives from the quality system supporting a State $PM_{2.5}$ monitoring network. Accordingly, the QPE process incorporates the principle of *graded approach* to provide wide flexibility in establishing the objectives and the measures.

The key to QPE is that management sets the performance objectives for the quality system based on the organization's mission and unique quality management needs. Experience has shown that externally-set objectives, such as those that might be set at the corporate level, may not be sufficient to adequately measure the quality system effectiveness for the quality system at a particular facility. By having managers set the performance objectives for a particular facility and mission, there is adequate flexibility in the process and the assurance that the performance measures reflect reality. As will be seen, this flexibility allows QPE to be applied to organizations of any size and quality systems of high or low complexity.

DESCRIPTION OF QPE

QPE is an ongoing internal management process and tool that uses indicators to convey information comparing an organization's past and present quality performance with its quality performance criteria. The process of QPE includes:

- selecting (e.g., developing and/or choosing) indicators for QPE;
- measuring (e.g., collecting data);
- analyzing and converting data into information describing the organization's quality performance;

- assessing information describing the organization's quality performance in comparison with the organization's quality performance criteria;
- reporting and communicating information describing the organization's quality performance; and
- reviewing and improving the QPE process.

QPE uses the concept of quality performance indicators (QPIs) in order to provide a measure of performance and to provide a basis for management evaluation. Two types of QPIs have been identified and categorized: management performance indicators (MPIs) and operational performance indicators (OPIs).

Management performance indicators (MPIs) are a type of QPI that provide information about management efforts to influence the quality performance of the organization's operations. MPIs relate to the policy, people, practices, procedures, decisions and actions at all levels of the organization.

Operational performance indicators (OPIs) are a type of QPI that provide information about quality performance of the operations of the organization. OPIs relate to:

- the design, operation, and maintenance of the organization's physical facilities and equipment;
- the materials, products, services, and wastes related to the organization's physical facilities and equipment; and
- the supply of materials, energy and services to, and the delivery of products and services from the organization's physical facilities and equipment.

QPE uses the Plan-Do-Check-Act concept that has been used extensively in quality management programs for more than sixty years, as shown in the following discussions.

Planning QPE (PLAN)

The identification of an organization's quality aspects is an important input in planning QPE. This information typically is developed in the context of a quality management system. An organization with a quality management system in place should evaluate its quality performance against its quality policy, objectives, targets and other quality performance criteria.

An organization should plan QPE in conjunction with setting its quality performance criteria so that the selected indicators for QPE will relate to and be appropriate for measuring or describing the organization's quality performance against the selected criteria. Examples of sources from which quality performance criteria could be derived include:

- past performance;
- customer or user requirements;
- best practices;
- management reviews and audits; and
- the views of potential customers and users.

In planning for QPE, management should also consider its organizational structure, overall business strategy, and quality costs and benefits. The financial, physical, and human resources needed to conduct QPE should be identified and provided by management. Over time, the scope of QPE can be changed to address other elements of an organization's activities, products and services that may impact the quality performance, as the organization changes to satisfy its evolving mission.

<u>Selecting indicators for quality performance evaluation</u>

Indicators for QPE help to condense relevant data into concise and useful information about

management's efforts, the quality performance of the organization's operations. An organization should select a sufficient number of relevant, significant, and understandable indicators to evaluate its quality performance. The number of selected indicators for QPE should reflect the nature and scale of the organization's operations. The choice of indicators for QPE will determine which data should be collected or which available data should be used. To avoid unnecessary effort, organizations may wish to use data already available and collected by the organization or by others.

Indicators for QPE are selected by organizations as a means of presenting quantitative or qualitative raw data or information in a more understandable and useful form. The information conveyed through indicators for QPE can be expressed as direct measures, or as relative, normalized or indexed information. Indicators for QPE may be aggregated or weighted as appropriate to the nature of the information and its intended use. Aggregation and weighting should be done with care to ensure that the data can be verified, consistent, comparable, and understandable. There should be a clear understanding of assumptions made in the handling of data and the transformation of the assumptions into information and indicators for QPE.

Moreover, any direct measurements of indicators must be supported by a sufficient and adequate system of quality control and quality assurance in order to assure that the indicator values are usable for QPE. Management may find that the application of techniques such as statistical process control (SPC) may be helpful in providing adequate quality in the indicators for QPE use.

There are many processes an organization may employ to select indicators for QPE, and several approaches that an organization may consider to select its OPIs and MPIs. Some quality aspects may be complex, and it may be beneficial to select a combination of QPIs to provide a comprehensive evaluation. Indicators for QPE should be selected so that management has sufficient information to evaluate the effect of progress toward achieving the quality performance criterion in one area has on performance in other areas of concern (e.g., synergistic effects).

Selecting management performance indicators

In the context of QPE, the management of the organization includes the policies, people, practices, and procedures at all levels of the organization, as well as the decisions and actions associated with the organization's quality aspects. Efforts and decisions undertaken by the management of the organization may affect the performance of the organization's operations, and therefore may contribute to the overall quality performance of the organization.

Management performance indicators (MPIs) should provide information on the organization's capability and efforts in managing matters such as training, customer and contractual requirements, resource allocation, documentation, and corrective action which have or can have an influence on the organization's quality performance. These MPIs should assist evaluation of efforts undertaken by management and actions to improve quality performance. For example, MPIs may be used to track:

- implementation and effectiveness of various quality plans or programs;
- efforts of particular importance to the successful quality management of the organization;
- quality management capabilities of the organization, including flexibility to cope with changing conditions, accomplishment of specific objectives, effective coordination, or problem-solving capacity; and
- compliance with contractual requirements and conformance with other requirements to which the organization subscribes.

In addition, effective MPIs may help to:

- predict changes in performance;
- identify root causes where actual performance exceeds or does not meet relevant quality performance criteria; and
- identify opportunities for preventive action.

Selecting operational performance indicators

OPIs should provide management with information on quality performance related to:

- the consumption of materials (e.g., processed, recycled, reworked, or raw materials; natural resources; energy; and services);
- the products (e.g., main products or by-products; recycled or reused wastes), services provided, and wastes produced (e.g., solid, liquid, hazardous, non-hazardous, recyclable, reusable), from the organization's operations that impact quality performance; and
- the physical facilities and equipment of the organization, their design, operation and maintenance, as well as the supply to and delivery from them, that impact quality performance.

Developing and using data and information (DO)

The following steps of the QPE process pertain to developing and using data:

- measuring (e.g., collecting data);
- analyzing and converting data into information describing the organization's quality performance;
- assessing information describing the organization's quality performance in comparison with the organization's quality performance criteria; and
- reporting and communicating information describing the organization's quality performance.

Collecting data

The organization should collect data regularly to provide input for calculating values for selected indicators for QPE. Any data should be collected systematically from appropriate sources at frequencies consistent with QPE planning. Data collection procedures should ensure data reliability and usability. This depends on factors such as availability, adequacy, and scientific and statistical validity. Moreover, the data should be verifiable. Data collection should be supported by sufficient and adequate quality control and quality assurance practices that ensure the data obtained are of the type and quality needed for QPE use. Data collection procedures should include the appropriate identification, filing, storage, retrieval, and disposition of data and information.

Analyzing and converting data

Data analysis converts collected data into information describing the organization's quality performance, expressed as indicators for QPE, which are useful for the organization's intended purpose. To avoid bias in the results, all relevant and reliable data that have been collected should be considered. Data analysis may include consideration of the data quality, validity, adequacy, and completeness necessary to produce reliable information. Information describing the organization's quality performance may be developed using calculations, best estimates, statistical methods, graphical techniques, or by indexing, aggregating or weighting.

Assessing information

QPE is intended to provide useful information on the management efforts of the organization and its

operations as a basis for appropriate management actions. The information, expressed in terms of QPIs, should be compared with the organization's quality performance criteria. This comparison may indicate progress or deficiencies in quality performance. The results of this comparison may be useful in understanding why the quality performance criteria have, or have not, been met. The information describing the organization's quality performance and the results of the comparison, should be reported to management, to support appropriate management actions to improve quality performance.

Reporting and communicating

QPE provides useful information describing the organization's quality performance to management for reporting and communicating to appropriate internal and external interested parties. Management should ensure that appropriate and necessary information describing the organization's quality performance is communicated throughout the organization on a timely basis. This may assist employees, contractors, and others related to the organization to fulfill their responsibilities, and the organization to meet its quality performance criteria.

Reviewing and improving quality performance evaluation (CHECK & ACT)

An organization's QPE process and results should be reviewed periodically to identify opportunities for improvement. Such a review may contribute to management actions to improve the quality performance of the management and operations of the organization, and may result in improvements in product and service quality. The success of the QPE process will be determined by the timeliness and effectiveness of the actions taken by management. QPE is a management tool. Unless it is used in a constructive manner, it will not yield effective results.

Management may also determine that the QPE process itself may need adjustment from time to time in order to keep the process current with other management and operational changes to the organization. Such adjustments will help to assure its continued effectiveness.

SUMMARY

QPE offers much promise as a process for enabling management to evaluate the effectiveness of a quality system in meeting the needs of the organization. Its successful application depends on managers setting realistic performance criteria and choosing performance indicators that represent the criteria. QPE may be applied to quality systems of almost any size and complexity.

The Appropriate Use of Professional Judgment in Sampling Design

Malcolm Bertoni, RTI , John Warren, USEPA, Kara Morgan, RTI

One of the fundamental choices a project team faces when planning to collect environmental data is whether to use judgmental sampling or random sampling. In random sampling (otherwise known as "statistical sampling" or "probability-based sampling"), the number of samples is determined through statistical calculations, and the selection of sampling locations involves some type of randomization process. In judgmental sampling, the number of samples and the selection of specific sampling locations are decided according to the judgment of the person(s) responsible for the sampling effort, often based on conditions observed in the field. Both approaches have their advantages and disadvantages. Advocates for judgmental sampling claim that it results in the most efficient use of time and money; skeptics point out that the resulting data carry severe limitations because they can't be used for quantifying uncertainty or estimating variability, and extrapolations cannot be made defensibly beyond the immediate area sampled. Advocates for random sampling claim that they provide a scientifically sound basis for drawing conclusions and quantifying the uncertainty in the results; skeptics point out that many statistical sampling designs require a large number of samples, often taken in areas that have no obvious value to the decision maker. Common misunderstandings and misconceptions of both approaches often lead to misapplications, which in turn can create a variety of quality, cost, and legal problems.

This workshop will help the participant understand how to decide when the use of judgmental sampling is appropriate and defensible, when the use of random sampling is most appropriate, and how professional judgment can be combined with random sampling to harness the advantages of both approaches. The session leaders will present a conceptual framework for understanding the two approaches, and participants will have an opportunity to discuss some recommended criteria for determining when judgmental versus random sampling should be used. The session leaders also will present some best practices for judgmental sampling, and identify strategies for incorporating professional judgment into random sampling designs, such that some of the cost and time advantages of judgmental sampling can be retained while maintaining the scientific defensibility of the sampling design, as well as the robustness of the resulting data.

WORKSHOP
Graded Approach Workshop for Assistance Agreements

Lou Blume, EPA, Quality Assurance Manager, Great Lakes National Program

The graded approach is advocated frequently in EPA QA requirements. The quality system documentation requirements for assistance agreements, such as grants are often cited as candidates for applying the graded approach, because many applicants have few resources for review and oversight and no direct Agency application for the results. Yet numerous quality guidance documents require quality management plans for environmental measurements and data generation and review. Thus, confusion and varied requirements exist regarding quality documentation for various types of assistance agreements. A cross agency workgroup was convened in FY2000 to attempt to address guidance examples for the varied types of agreements.

This workgroup conducted a well attended fact finding session during last years National Conference on Managing Environmental Quality Systems and has continued to meet via conference. Using an outline for a guidance report that follows the presentation format below, each of the presenters will be sharing examples along with recommended graded approaches for developing quality system documentation. The workgroup will convene to an ad-hoc writing session to complete the first draft of a guidance document graded approaches for assistance agreements.

A Graded Approach to Documenting the Use of Existing Data by Assistance Agreement Holders

Patricia Lafornara, EPA

This is an illustration of ways a graded approach might be applied to documenting project uses for existing data in assistance agreements. Also included is a preview of guidance being prepared for the use of existing data in a revision of the Guidance for Quality Assurance Project Plans (EPA QA/G-5).

This presentation describes guidance for a graded approach to documenting using existing data – not for an approach to determining whether the data are of sufficient quality to use. It is presumed that you determined whether the data are adequate for your intended use, or plan to do so as part of your project. Please note, too, that if your data are inadequate, you must report any limitations and how this affected your project.

The elements of the quality assurance project plan (QAPP) that involve the use of existing data are B9 and perhaps B10, C1 and C2, and D1, D2, and D3. Let's presume that the A elements for project management would apply to every project. Therefore, this is the boundary for everyone's quest for a maximum set of QAPP elements for a project involving ONLY existing data! This is not a graded approach, but rather selecting elements that apply. Technically, elements that don't apply should be listed with the appropriate rationale for not including them.

What's special or unusual about existing data, however, is that they can be a part of a project at more than one phase in the project life, and they may be quantitative or qualitative. Existing data can be used, for example, to scope the project; for some or all of data input to a model, GIS project, or a risk assessment; to substitute for some new data collection in a sampling-based project; and/or for evaluating project results. Again, this is not a true issue for a graded approach, just an issue with documentation oriented to new sample collection and analysis.

For any application, the intended use is the most important consideration, but resource constraints follows closely. For example of a scoping phase application of existing data, imagine selecting an unapproved method for low-level analysis. If a method will be selected based upon a search of results reported by others (in journal articles, perhaps) rather than by the project personnel themselves, it may be important to report the search criteria and scope of the search. If the method selected is critical to the project, information about the selection process for verification would be helpful. For documentation, say what the information or data were and where they came from (how they were discovered, perhaps), their importance to the project, and the acceptance criteria considered, if any. These points make up element B9. If resources are a problem, listing that a method will need to be chosen may be sufficient. Perhaps confirmation of a selection method can be done informally. In this example, B10 will not likely apply.

Consider a different example, perhaps of researching a toxicological value for a risk assessment project, however. The literature searched will be important, and a system for handling/sifting and recording the possible studies (and data from them) may be necessary. This data management activity would be included in B10, as it goes beyond just stating acceptance criteria for the candidate studies. If resources are limited, a simple search of a government data base may be all that is anticipated, however, and one line may suffice to alert QAPP reviewers to the limiting search terms planned. Resource limitations are obviously going to affect more than the documentation aspect of a project – if you don't fund a thorough project, you will have a similar quality result. If you are looking for a solid quality assurance approach, Doug Fennell in NCEA (ORD/EPA), authored a very careful work on compiling searches for literature on research to document and interpret risks of pollution to human health and the environment. He's hoping

to publish it, but you can contact him for a copy by asking for "A White Paper on Compiling and Evaluating Research Information for Secondary Use in ORD's Assessment Documents."

In some trend-seeking projects, like those based upon a major search for historical data, the QAPP C elements may actually comprise the project. In others, they may play a smaller role. When using existing data to scope a project, for example, the assessment element will usually apply to the planned project rather than the data used in scoping. Imagine that you have a project scoping exercise using qualitative information like fish species reported to be in a series of lakes (the project will use a subset of these lakes). Assessing the use of that species information may be unnecessary if it is in the context of one of many lake selection criteria. However, if it's critical to the project success to identify whether certain fish species are present or absent, a careful assessment of how these data are used may be prudent! This would be described in C.

The use of the D elements is related to the nature of the existing data. What "pedigree" accompanies the data, or whether a "pedigree" is important to the project are both considerations. Again, if resources limit the extent of verification and validation planned, be aware of the effect on the quality of the final result. *Guidance on Data Quality Assessment* (EPA QA/G-9) is recommended as a source for good practices.

In summary, there are several QAPP elements that are appropriate for documenting existing data use. Applying the graded approach to QAPP documentation is quite distinct from applying the graded approach to planning a project without adequate consideration of existing data used. The latter is not recommended! You are encouraged to show care in documenting aspects of planned use that are important to obtaining satisfactory project results. Below is a partial, draft table of selected QAPP elements. It includes some considerations that may be addressed within the elements concerning the identification and use of existing data. The table is being considered as part of the revised QAPP guidance. The most recent text proposed for element B9 will be distributed at the work shop.

GROUP B: DATA GENERATION AND ACQUISITION	
B1: Sampling Process Design (Experimental Design) B2: Sampling Methods B3: Sample Handling and Custody B4: Analytical Methods B5: Quality Control B6: Instrument/Equipment Testing, Inspection, and Maintenance B7: Instrument/Equipment Calibration and Frequency B8: Inspection/Acceptance for Supplies and Consumables	Elements B1 through B8 address various quality aspects of the design and procedures for collecting, handling, and analyzing environmental field samples. They are generally relevant only when collecting new data for purposes of addressing the project's objectives. Thus, these elements generally do not address issues regarding existing data sources. In some cases (for example, on projects using exclusively existing data), the project's principal investigator may decide to present certain procedures associated with the generation and use of existing data within these QA Project Plan elements rather than all appearing in element B9. However, it is often cleaner to have elements B1 through B8 focus only on newly-generated data and to have element B9 focus on existing data.
B9: Non-Direct Measurements	This is the primary element of the QA Project Plan within which information on existing data, their intended uses, and their limitations is presented. This section also presents the acceptance criteria for specific data sources that may have been introduced in element A7.
B10: Data Management	This section documents how existing data (as well as newly-generated data) would be incorporated and managed into the project's data management system. Example topics include how existing data will be obtained from its source in a given format, how and what data will be entered and verified if obtained in hard copy format, and how certain security or confidentiality requirements will be incorporated into the project's data management system.

GROUP C: ASSESSMENT AND OVERSIGHT	
C1: Assessments and Response Actions	Examples of assessments that involve the use of existing data that may be implemented in the project (and thus documented in this section) are the following: C Assessments that existing data meet basic project requirements (e.g., are of the proper type) and are appropriately relevant and suitable for their targeted use (e.g., has an acceptable target population). C Assessments that the quality of existing data meet the acceptance criteria specified in Sections A7 and B9 and that a sufficient quantity of existing data is available to allow the project to meet criteria on data quality. C Assessments that proper procedures and protocols were used in obtaining or abstracting existing data from their sources. C Assessments that sufficient quality control information were obtained on these data. C Assessments that the quality assurance techniques documented in the QA Project Plan have been followed in the use of the existing data. Assessments involving existing data generally address the process of acquiring, evaluating, selecting, and obtaining existing data for use on the project, and then using the data in the manner in which the data were considered acceptable. A graded approach is used to determine the overall scope and level of detail in which the assessments are performed. The following types of information on these assessments should also be included in this section (as they would be for any type of assessment): C The role that these assessments play in the project's total set of assessments C The schedule of assessments C The organizations and individuals expected to participate in the assessments

	C Information expected from the assessment C Documentation requirements for the assessment C Possible types of corrective action and levels of authority that would determine corrective action (e.g., collect additional data, investigate other data sources, loosen acceptance criteria).
C2: Reports to Management	Cite any reports that need to be brought to the attention of management that may affect the extent to which the project relies on existing data.

GROUP D: DATA VALIDATION AND USABILITY	
D1: Data Review, Verification and Validation	This section would document how the ability to use existing data to achieve the project's requirements will be evaluated. While the assessments in element C1 may have initially been performed on existing data, this section discusses the final set of assessments of how the data can be used to address project objectives. Although previous sections of the QA Project Plan address how an entire existing data source is determined to be acceptable for use on the project, this section would address how individual data values and information within the existing data source are determined to be acceptable for use or otherwise need to be qualified, when the procedures would be performed, and by whom.
D2: Verification and Validation Methods	Any mathematical or statistical procedures (such as outlier analyses or goodness-of-fit tests) that will identify whether individual data values within existing data sets should be rejected, transformed, or otherwise qualified prior to any statistical analysis would be discussed here. In addition, if existing data need to be entered into a project database, the features of the data management system that verify the accurate entry of values for important data parameters into this database, along with any data reduction procedures (e.g., averages of replicate measurements), will be detailed here. The point in the schedule at which these activities will be done should be mentioned.
D3: Reconciliation with User Requirements	The ultimate "adequacy" of the existing data in this project relative to the data users' requirements is determined by methods detailed in this section. This is done by describing statistical tools and other methods used to evaluate whether the existing data can be used to achieve their intended uses and are therefore justified to be used in addressing project objectives. Such statistical tools are documented in *Guidance for Data Quality Assessment: Practical Methods for Data Analysis (EPA QA/G-9), QA00 Version* (EPA, 2000). Strategies in place to resolve or account for any issues that arise from investigating the data (e.g., impact of data limitations that were encountered, need for new data collection, re-analysis, use with caveats) would be discussed.

Workshop on
Improving the Quality System Specifications for EPA's Contracts

Brenda Young, EPA Quality Staff
Al Batterman, National Health and Environmental Effects Research Laboratory
Mark Doehnert, EPA Office of Radiation and Indoor Air
Holly Ferguson, EPA National Risk Management Research Laboratory
Nan Parry, National Center for Environmental Research
Ann Vega, EPA National Risk Management Research Laboratory

EPA is currently revising its quality-related policies for solicitations and contracts. This is a result of changes to the Federal Acquisition Regulations (FAR), new flexibility in contracting procedures such as simplified acquisition, and new types of contracting such as performance based contracts. A workgroup of quality assurance (QA) professionals from across the Agency have been charged with updating the Agency's quality-related policies for contracts and solicitations to address these changes and take advantage of the new flexibility and procedures. This session will describe current changes to EPA's internal policies, provide an application of these policies to an EPA program, and discuss other contracting issues that will result in further changes. The workshop will close with a panel discussion on simplified acquisitions in which workshop participants will be encouraged to share their experiences, raise issues, and make recommendations for changes.

The FAR 46.202 describes the four categories of contract quality requirements, depending on the extent of quality assurance needed by the Government for the acquisition involved. A 1999 revision to FAR 46.202-4, *Higher-level Contract Quality Requirements,* and the corresponding contract clause at 52.256-11 now allow Federal agencies to select a national consensus standard as the basis for their higher-level quality requirements for contracts. The FAR contract clause also allows tailoring of the standard to more effectively address specific agency needs or purposes. Based on these FAR provisions, the Environmental Protection Agency (EPA) has selected ANSI/ASQC E4, *Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs*, as the basis for its higher-level environmental quality requirements, and, as permitted by the FAR, will tailor this standard to ensure that Agency needs are met.

Due to these changes in the FAR, EPA Acquisition Regulation (EPAAR) 1546.2, *Contract Quality Requirements* (Mar 1984), which was a quality regulation that applied only to EPA, became unnecessary. On March 20, 2001, this EPA-specific regulation was removed along with its clauses (1552.246-70, 71, and 72) through a Direct Final Rule that was published in the Federal Register (65 FR 79781, December 20, 2000). The tailoring language allowed by FAR 52.246-11 and pertinent requirements from EPAAR 1546.2 will be incorporated into a revision to EPA Directive 1900, *Contracts Management Manual* (CMM). A procurement policy notice (PPN 01-02) was issued March 20, 2001 to ensure an orderly transition from EPAAR 1546.2 to the CMM. It contains sample clauses which include the tailoring language to the ANSI/ASQC E4 standard, as allowed by FAR 52.246-11. PPN 01-02 allows greater flexibility in meeting EPA's documentation requirements.

Measurement Uncertainty Expression
of
Sampling and Testing Data

Marlene Moore, Advanced Systems, Inc.

The presentation is a practical approach to measurement uncertainty using the existing data generated as part of the routine quality control and quality assurance practices defined in the ANSI/ASQC E-4 standard for data collection activities. The expression of the results with the measurement uncertainty provides the final data user with a uniform method of comparison environmental data *of the data quality.*

The presentation demonstrates the use of the international definition for expressing measurement uncertainty when reporting values for environmental compliance and site investigations or clean-up. The presentation combines the approaches from various authors including the Corp of Engineers and U. S. Navy with the EPA approach defined in the R and G series documents published by the Quality Assurance Division.

The ANSI/ASQC E-4 defines the elements required for a quality system that is involved in any collection activity. The use of the nested approach provides an objective measure of the performance of the quality system for both the data user and the management of the quality program for each project.

The nested approach, developed by Mr. William Ingersoll, U.S. Navy, NAVSEA, presents the data using the measurement uncertainty as defined in the international standard and the correction for method bias as presented by the U.S. Army Corp of Engineers in several recent papers. The nested approach provides a uniform and objective process for reporting the variance in the sampling and testing data used for all environmental regulatory programs. This presentation assumes that the sampling and testing planning, implementation and validation incorporate a mature quality system approach to minimize or eliminate mistakes.

The international definition of measurement uncertainty has been defined in the "Guidelines for Expression of Measurement Uncertainty" (GUM). This uniform definition requires the reevaluation of the uncertainty expressions being used when expressing measurement results. The expression of measurement uncertainty is required when reporting a value and is not applicable to qualitative methods, such as presence/absence or positive/negative result reporting.

In current environmental programs, measurement uncertainty is expressed for instrumental uncertainty such as radiochemistry, and for probability distributions such as biological testing. However current environmental programs do not define the reporting of measurement uncertainty. Measurement uncertainty must include sampling, site characteristics, matrix effects and the laboratory effects. Therefore the definition and determination of measurement uncertainty is required by the data user and is not a laboratory-generated value. The laboratory provides the data for calculating the uncertainty, but the data user must generate and define the measurement uncertainty for environmental reporting and decisions.

The uncertainty provides the data user with the interval about the result. This interval expresses the random and systematic effects on the measurement. The values present the variability of the result, thus providing a level of confidence when making a decision using the result. With the adoption of a single

standard, (GUM) for defining uncertainty for the international presentation of data, the use of the plus/minus symbol should become more commonplace in the future.

Uncertainty expresses the range of values that could reasonably be attributed to the measured quantity. The expanded uncertainty provides the level of confidence that the value actually lies within the range defined by the uncertainty interval.

The estimating of uncertainty is a quantitative indication of the quality of the result. This estimation provides the data user with the confidence to allow comparability. This is needed in order to, allow accreditation bodies an objective approach to resolving data comparability complaints, provide the data user with information related to the risk in making a decision and provide the regulatory with the variance associated with the data when comparing to risk or other type of regulatory criteria.

An example may help to explain this idea. The nested approach estimates the measurement uncertainty of a value. The uncertainty is estimated for copper in wastewater using data generated as required by the Clean Water Act under the National Pollution Discharge Elimination System (NPDES) program. The samples and related quality control data are routinely collected and are represented in Table 1 as the following:

IME - Intrinsic instrumental measurement effects

SPE - Spike preparation effects

PME - Preparation method effects

MIE - Matrix interference effects

SCE- Sample collection effects

SLE - Sample location effects

From this example sample collection effects and sample location effects are not apparent due to the lack of field duplicate samples collected and collocated sample collection information. If the sampling design for these wastewater events included the collection of a duplicate field sample and collocated sample then the uncertainty for the entire measurement may be estimated. A collocated sample may be taken using a second composite sampler over the same time frame. Since the sampling tube would be placed in the same stream, but a different part of the stream, the sample would be collocated and provide information on the heterogeneity of the wastewater stream.

Using the information as we commonly see collected for wastewater in this example, we can see that if the action level or permit limit is 5.0 mg/L, the uncertainty expression indicates that there is a probability that the wastewater may be above this limit at times. This contradicts using just the average result, which would indicate that the wastewater is always below the limit.

This uncertainty expression provides the data user with specific ranges for both the measurement uncertainty with out error correction (bias) and with error correction. This allows the data user performing risk or other types of data assessment with the effect of the bias on the measurement system. The reporting of the error corrected values would not be necessary if the method used is mandated by the EPA regulation and a regulatory limit has been established for the specific program using that method. However as EPA continues to allow method flexibility the need for a sound basis of comparison is required. This basis must address not only laboratory data, but also the sample type, sampling method and site variability.

All measurement operations use quality assurance statements so the structure of the variance in the data is

defined and known. These quality assurance statements have been identified in the past as the PARCC's or precision, accuracy, representativeness, comparability and completeness. New terms are emerging such as measurement quality objectives (MQO) and data quality indicators (DQI). These terms are often misapplied or interpreted due to the various regulatory and other program definitions. In all cases however, each term is used to categorize the measurement variance into different groups with the expectation that the data user will interpret these sources of variance based on the regulatory or data needs.

In many cases, users of data attempt to make it appear that the variation is a result of poor sampling or poor measurement procedures, when in fact; it is a natural scientific phenomenon, when all these procedures are not designed for the specific activity. Some data users assume that if reference methods are followed the variation does not affect the decision making process no matter how far a field the application is to the original procedure design. For example, wastewater methods are often used for measuring fish tissue or biosolids when the original method studies were based on the testing of waters with less than 1000 ppm solids. (ppm = parts per million)

The sampling and testing design, implementation and validation must address this variation. Every aspect of compliance monitoring and site investigations and remediation activities must know the sources of these variations in order to minimize the effect on the data used for decision making. The demonstration of an understanding and knowledge of the measurement process ensures the generation of defensible data. The use of the nested approach calculator provides the data user with a graphical and numerical presentation of the data to identify the significant components of the uncertainty. This allows planners and data uses to focus on the significant contributors to the uncertainty found in the data. The identification of the significant sources contributing to the uncertainty ensure that those aspects of the sampling and testing operations are optimized to minimize and reduce the uncertainty in the results.

Currently, many compliance and data generation applications require the use of data qualifiers or letter codes to provide the data user with some information related to the data acceptability. These codes are not uniformly applied and are often confusing since they do not provide objective information. These codes often present subjective information such as estimated results, or blank contamination. It is often not possible to determine the reason the code was applied to the number, such as unacceptable calibration criteria, unacceptable surrogates performance or others. The criteria used differ by the group applying the code. This has lead to confusion and rejection of data that may be usable when the variance of the data is more uniformly presented.

The measurement uncertainty principles allow a single uniform expression of the measurement based on international definitions. The quantification of the uncertainty using the nested approach is the first step. After the quantification identifies the significant components of uncertainty, the project manager must detail and better define these components to ensure that the variance from these components are minimized.

References:

"Guide to the Expression of Uncertainty in Measurement", ISO, Geneva, Switzerland 1993 (ISBN 92-67-10188-9) Known as "GUM"

"International Vocabulary of Basic and General Terms in Metrology", VIM ISO(1993), second ed, International Organization for Standardization, Geneva, Switzerland, ISBN 92-67-01075-1

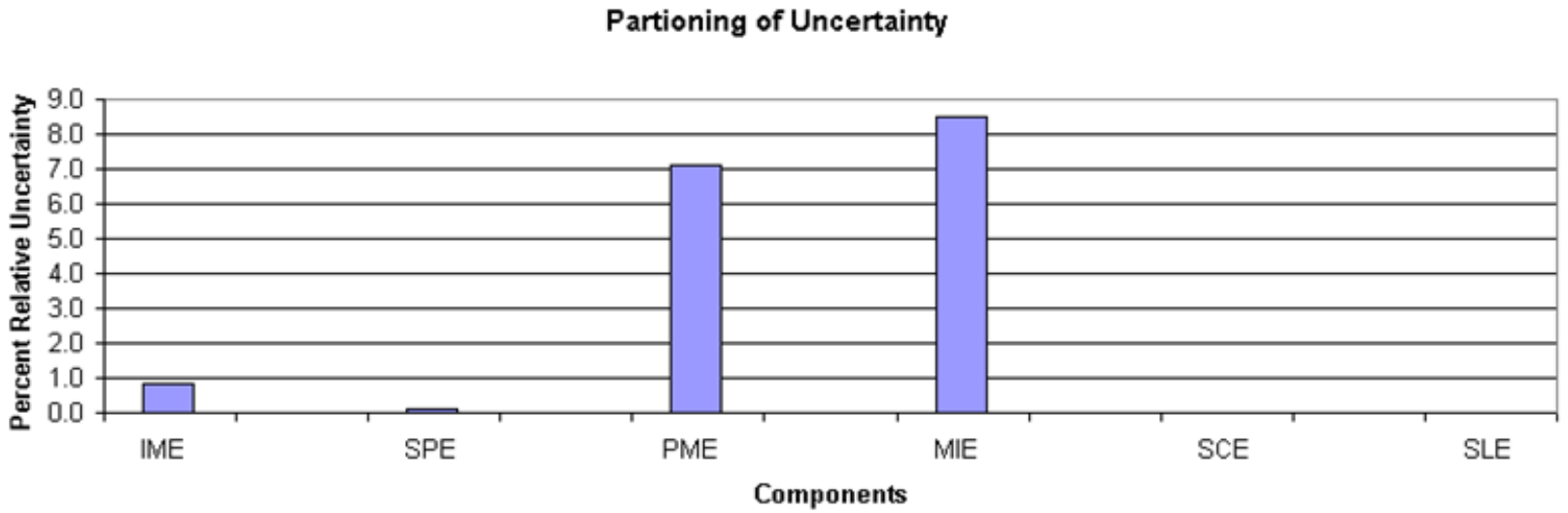
"Environmental Analytical Measurement Uncertainty Estimation: Nested Hierarchical Approach", U.S. Navy DTIC, #ADA396946, William Ingersoll, Charleston, S.C, 2001

"Estimation of Laboratory Analytical Uncertainty Using Laboratory Control Samples", Thomas Georgian, U.S. Army Corps of Engineers, "Environmental Testing & Analysis", Nov/Dec, 2000 pp 20-24, 51

“Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs.” American National Standard, ANSI/ASQC E4-1994

Table 1

Copper in Wastewater = Analytical Measurement Uncertainty Calculator



What is the measurement result?	**4.5**
What are the measurement units?	**mg/L**

If the sample measurement is 4.5 mg/L then the uncertainty interval is 3.5 - 5.5 mg/L at the specified Confidence Level Expanded Uncertainty.
For the above result, if the systematic measurement error (bias) is corrected, then the uncertainty interval is 3.3 - 5.3 mg/L at the specified Confidence Level Expanded Uncertainty.

The information presented in this table is supplied by Mr. William Ingersoll, NAVSEA

WORKSHOP
$PM_{2.5}$ Ambient Air Monitoring Program: Use of Data Quality Objective Software tool for Data Quality Assessments

Michael Papp, U.S. EPA

Data quality objectives (DQOs) are qualitative and quantitative statements derived from the DQO Process that clarify the monitoring objectives, define the appropriate type of data, and specify the tolerable levels of measurement errors for the monitoring program. By applying the DQO Process to the development of a quality system for $PM_{2.5}$, the EPA guards against committing resources to data collection efforts that do not support a defensible air quality management program. One of the primary objectives for collecting ambient air data from the State and local Ambient Monitoring Site (SLAMS) network is for comparison to National Ambient Air Quality Standards (NAAQS) The DQO Process discussed at the Workshop illustrates the steps taken to assess the quality of data needed for making comparisons to the $PM_{2.5}$ NAAQS and describes the use of a DQO software tool for determining the "gray zones" for both the annual and daily NAAQS. The tool can be used to help individual State, Local and Tribal monitoring organizations assess whether their quality systems are in control. The workshop will also cover two additional issues related to PM2.5. OAQPS will detail the specific procedures to be used in computing PM2.5 NAAQS 'design values', with a focus on data completeness issues. Some time will be devoted to discussing the QC/QA aspects of the recently released re-engineered Air Quality System [formerly part of the Aerometric Information and Retrieval System (AIRS)]

DQOs for $PM_{2.5}$ were developed during the months from April to July of 1997. A number of assumptions were made in order to generate realistic error rates. Table 1 provides a listing of these assumptions. In 2001, EPA reassessed the assumptions underlying the 1997 DQOs. In almost all cases, the assumptions made in the 1997 process held true in the 2001 evaluation.

Table 1. 2001 DQO Assumptions

1. Bias is -10% or + 10%
2. Precision is 10%
3. Annual NAAQS is controlling standard
4. No spatial uncertainty and each monitor stands on its own (no spatial averaging)
5. 1 in 6 sampling with 75% completeness (144 days)
6. 3-year annual average is truth, (every day sampling and 100% comp.) up to bias and measurement variability
7. Log normal distribution for population variability, 80% CV
8. Normal distribution for measurement uncertainty
9. Seasonal ratio (ratio of avg conc for highest season to lowest season) = 5.3
10. No auto correlation in daily concentrations
11. Bias and measurement variability (precision) applies to entire 3 years
12. Type I and type II decision errors set to 5%

The $PM_{2.5}$ DQOs were generated using conservative but realistic assumptions. For example, the DQOs were generated assuming a sampling frequency of every 6 days with 75% completeness. This is the lowest sampling frequency allowed in the Code of Federal Regulation. A 95% confidence limit around the annual mean at this sampling frequency would be "wider" than a 95% confidence limit for an every day sampling frequency at 90% completeness. In all cases, the assumptions in Table 6-1 are close to the

extremes of realistic and allowable data. Assumptions in bold are variables that will be used in the DQO Software tool. Figure 1 provides the power curve for the three year annual mean based on the 2001 assumptions shown in Table 1. A power curve is an easy way to display the potential of decision errors based upon the choice of various assumptions that affect data uncertainty. The gray zone is the range of concentrations for which the decision errors are larger than the desired rate of 5%.

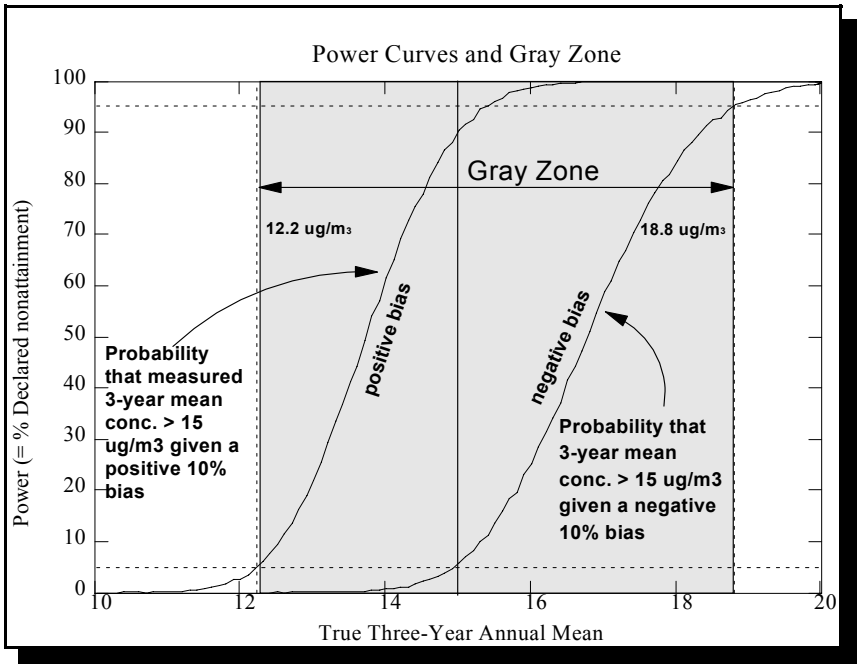


Figure 1. Power curve based on 2001 assumptions

Based on the 2001 assumptions, the gray zone is 12.2 to 18.8 Fg/m^3. This means that if all the 2001 assumptions hold, the decision maker has a 5% chance of observing a 3-year mean concentration that is greater than 15 Fg/m^3 even though the true mean concentration is 12.2 Fg/m^3. As has been mentioned, the 2001 assumptions are realistic but conservative. For example the CY00 $PM_{2.5}$ QA Report demonstrates that the precision and bias estimates at a national level are well within the DQOs. Assumptions that are "better" than those listed in Table 1 will tend to decrease the width of the gray zone. Figure 2 provides an example of the power curve/gray zone changes for a simple change in sampling frequency from 1 in 6 day (green/solid) to 1 in 3 day (blue/dots) to every day (red/dashed); all the other 2001 assumptions remain the same. Higher sampling frequencies result in narrower gray zones, meaning that decision errors are reduced. Because there is potential for the assumptions to vary, OAQPS commissioned the development of a software tool to help Headquarters and State, local and Tribal organizations determine the potential for decision errors based on assumptions relevant for sites within their network. Figure 2 is generated using this tool and allows for multiple scenarios (power curves) to be reviewed on one table. The assumptions listed in bold in Table 1 can be changed to suit a particular network.

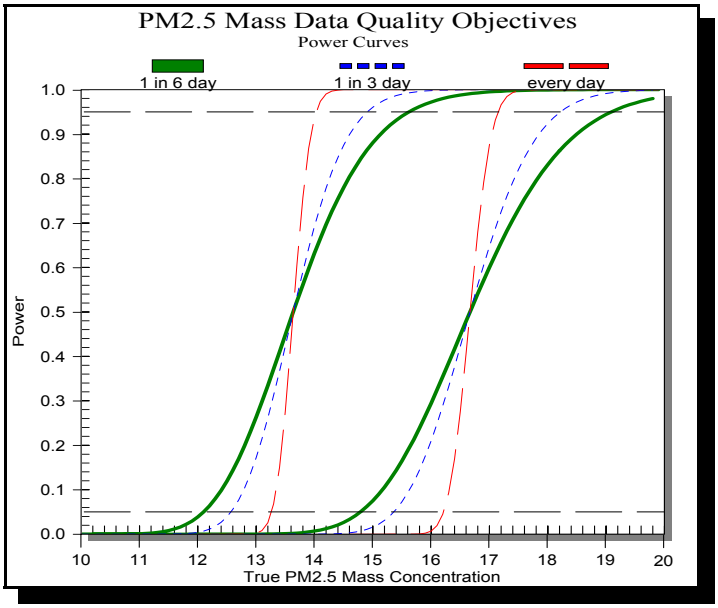


Figure 2. Power curve changes due to changes in sampling frequency

The DQO evaluation showed that sampling frequency, population variability (assumed to be lognormally distributed with a CV of 80%), and measurement bias play a significant role in the width of the gray zone. Measurement precision did not have a significant effect on the gray zone which suggests more imprecision could be tolerated with little effect on decision errors (when evaluating an annual mean developed with 3 years of data). The $PM_{2.5}$ mass DQOs were developed for making good decisions about the 3-year average of annual means, since it was assumed that the annual standard was the controlling standard. In particular, they were developed to evaluate the chance of concluding an average concentration was above 15 $\mu g/m^3$, when in truth it was not, and the chance of concluding an average concentration was below 15 $\mu g/m^3$, when in truth it was not. Due to the number of measurements that go into the 3-year average of annual means (at least 144), it is easy to see why measurement precision does not have a large influence on the size of the gray zone of the power curve. If, however, the DQO tool displayed the power curves for the daily standard (the 3-year average of the annual 98th percentiles), it is likely that measurement precision would be important for the decision errors, since the extremes of distributions are less robust than the centers. Recent evaluations of the continuous monitors have shown precision estimates comparable to the FRMs.

Data uses that involve no averaging, such as real-time reporting, are even more sensitive to measurement imprecision. Thus, caution should be exercised in drawing conclusions from the DQO power-curve tool. The tool has been designed for specific data uses, namely, evaluating decision errors associated with the $PM_{2.5}$ standards and is based on specific assumptions. If the assumptions are not appropriate or if the data use is different than comparison to the standards, the power curves and gray zones likely do not reflect the true decision errors.

Example

The DQO tool is being enhanced to present both forms of the standard so that is can be used as a data quality assessment tool for State, Local and Tribal agencies. to ensure that decision errors are acceptable for both standards. Figure 3 provides an example of the power curve for a 3-year mean based on the

following data quality input parameters

1.	bias	12%
2.	completeness	90%
3.	sampling frequency	1 in 3 day
4.	measurement CV	10%
5.	population CV	50%
6.	Seasonal ratio	5.3

Even though the bias was higher that the 10% acceptance criteria, because the state had more complete data and a higher sampling frequency than 1 in 6 day, the gray zone is "tighter than the annual mean DQO. Therefore, in this example, the State monitoring network meets the PM2.5 annual standard DQO.

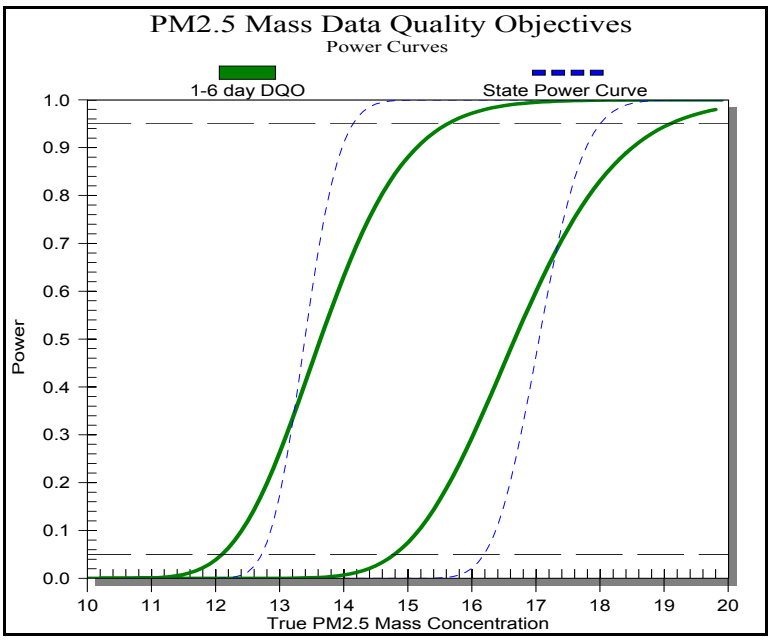


Figure 3. Annual mean power curve compared to State power curve

Preparation of a Standard Operating Procedure for an Emergency Quality Assurance Project Plan (Qapp)

Christian Byrne, US. EPA, OPP/BEAD/ECL
Betsy Grim, QAM, OPP

Following the events of September 11, 2001 in New York City, New York and Washington, D.C., a great amount of environmental sampling was undertaken to determine the level and extent of toxic material contamination at these sites. United States military units as well as Federal civilian agencies were requested by the President to assist in these activities. As has often been the case in unexpected, catastrophic events, there was often a lack of coordination and communication between and among the organizations requesting assistance, organizations tasked to begin the assessment of the event and arrangement of sampling of the site, and the organizations tasked to perform the analysis of the samples taken. In several instances that ECL was involved in, samples were taken with no discussion of the sampling protocol interface between organizations; no clear purposes were discussed as to the levels of detection needed or use of the resulting data to facilitate acknowledgement of excessively high ("hot") concentrations; or the defined chain of command for communication was not made available. It is understood that in these types of "unique" circumstances, there is little time to formally prepare a quality assurance project plan (QAPP) or, even, a small project quality assurance project plan (spQAP). However, the situation exists for poor planning, inexact sampling, and the generation of data which eventually is judged to be unacceptable, resulting in wasted time and resource allocation. In these circumstances, this could result in injury and/or death. Actions or lack of properly conceived actions might result in ligation from the affected parties and questions regarding the lack of adherence to quality requirements by the Agency.

The establishment of a standard operating procedure for the preparation of an emergency quality assurance plan (EQAPP) should be considered and reviewed. This emergency QAPP would encompass one (1) page with the most basic points of contact, planned purpose by the requestor, established sampling protocol, analytical method, and final report. This would require some degree of time to arrive at this information and later a more comprehensive QAPP would be prepared to document all of the elements finalized. It is extremely important to have a turn-key document as this in place when this type of event occurs and perceived effort and emotions overcome proper planning and objectives.

Sequential and Adaptive Sampling Approaches Within Visual Sample Plan (VSP) In Support of Dynamic Field Activities

Brent Pulsipher, John Wilson, Dick Gilbert, Nancy Hassig-PNNL

With more field analytical devices becoming available, real-time, in-field decisions are more feasible. Guidance is forthcoming on Dynamic Field Activities that combine on-site data generation with on-site decision-making [1]. Drawbacks of traditional characterization campaigns include the delay between data gathering and decision- making and the difficulty of returning to the field for more samples if the desired Data Quality Objectives (DQOs) are not met [2]. Use of in-field analytical methods can significantly streamline the characterization and decision-making process by permitting real-time decisions. Figure 1 contrasts the traditional characterization approach with the dynamic field activities approach. Note that the dynamic approach does not circumvent the need for systematic planning and DQO development.

Statistical Methods and Tools

Many of the classical statistical sampling schemes and available tools for determining the number of samples required for confident decisions are most applicable when all samples are obtained and analyzed before decisions are made. However, with field analysis capabilities available, one may want to employ a sequential or adaptive sampling approach. Statistical methods exist for sequential and adaptive sampling but they have not been readily available to the non-statistician in the form of software tools. Under these sequential sampling approaches, additional samples are collected until a confident decision (e.g., clean vs. contaminated) can be made based on data that are available, whereupon sampling can stop. It has been shown that statistical sequential sampling approaches result in fewer overall required samples than the traditional single sampling campaign approach under most conditions.

Visual Sample Plan (VSP), a software tool being developed with support from DOE, EPA, and DoD, facilitates the quantitative aspects of the Data Quality Objectives (DQO) process goals of obtaining the right type, quantity, and quality data to support confident decisions [3]. Two sequential sampling modules and an adaptive sampling module are available within VSP that support not only the single sampling campaign but also a dynamic sampling plan. This tool will be demonstrated for the case when field collection and analysis devices are available.

VSP Sequential Sampling Design Module

Sequential methods are iterative in nature such that some data are gathered and an evaluation is performed to determine whether sufficient information exists to support confident decisions. If so, then no further characterization is required. If not, then additional characterization data is obtained via more samples/analyses. This process of data gathering and assessment is continued until either a confident decision is supported or it becomes no longer cost-effective or feasible to continue.

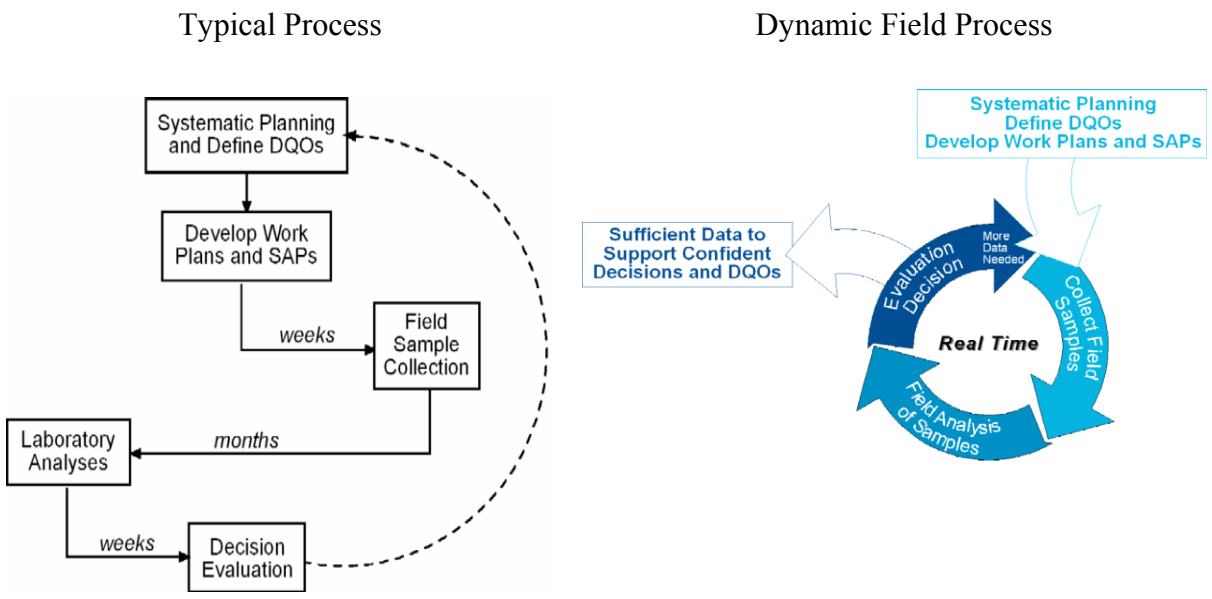


Figure 1. Typical vs. Dynamic Field Process

Two sequential sampling approaches are provided within VSP. For each of these VSP modules, the primary sampling goal is to compare the mean concentration against a regulatory threshold. Sequential methods for other sampling (decision) goals are not yet available within VSP. The Sequential Probability Ratio Test module is appropriate when the sampling and analytical standard deviation is previously known, perhaps from numerous previous studies [4]. For most cases, the sampling and analytical standard deviation is unknown and will be estimated from the gathered data. In that case, the Barnard's Sequential Test module is appropriate [5]. The specific statistical methods will not be detailed in this report but will be available with documentation that will accompany a future release of VSP. Although the underlying calculations are different for these two modules, the VSP dialog boxes and outputs are similar.

The sequential modules require the following DQO input as shown in the VSP dialog box in Figure 2.

- Whether site is assumed dirty until proven clean or clean until proven dirty.
- Acceptable probability of concluding dirty if truly clean.
- Acceptable probability of concluding clean if truly dirty.
- Width of the gray region or "delta".
- Number of samples/analyses obtained per sampling campaign (trip to the field).
- Sampling and Analytical Costs.
- Site map or sampling area of concern.
- Action level or regulatory threshold.
- Known or estimated sampling and analytical standard deviation.

Because an initial estimate of the sampling/analytical standard deviation is required for the Barnard's test, VSP requires that at least 10 samples are obtained and the contaminant concentration results entered into VSP. Then VSP performs the statistical test to determine whether the data supports a decision that meets the required DQOs. If so, then VSP recommends a particular decision without any further sampling. If not, because this is an iterative procedure, VSP will recommend that more samples should be obtained and

in-field analyses be performed with the resulting concentration data to be input into VSP. Specific output from VSP at each step of the iteration, as shown in Figures 2 and 3, includes

- Evaluation of whether a confident decision given DQOs is supported.
- Recommended sampling location(s) (x,y coordinates) at each step of the iteration depicted on a site map.
- Projected number of samples that may have to be obtained to support confident decision.
- Visual graphic depicting the mean concentration at each iteration and the decision recommendation.
- Calculated mean and standard deviation.

Details on the use of these methods including how to input data, interpretation of VSP output, limitations of methods and software, creation of maps, and other useful information will be provided during the presentation at the EPA meetings.

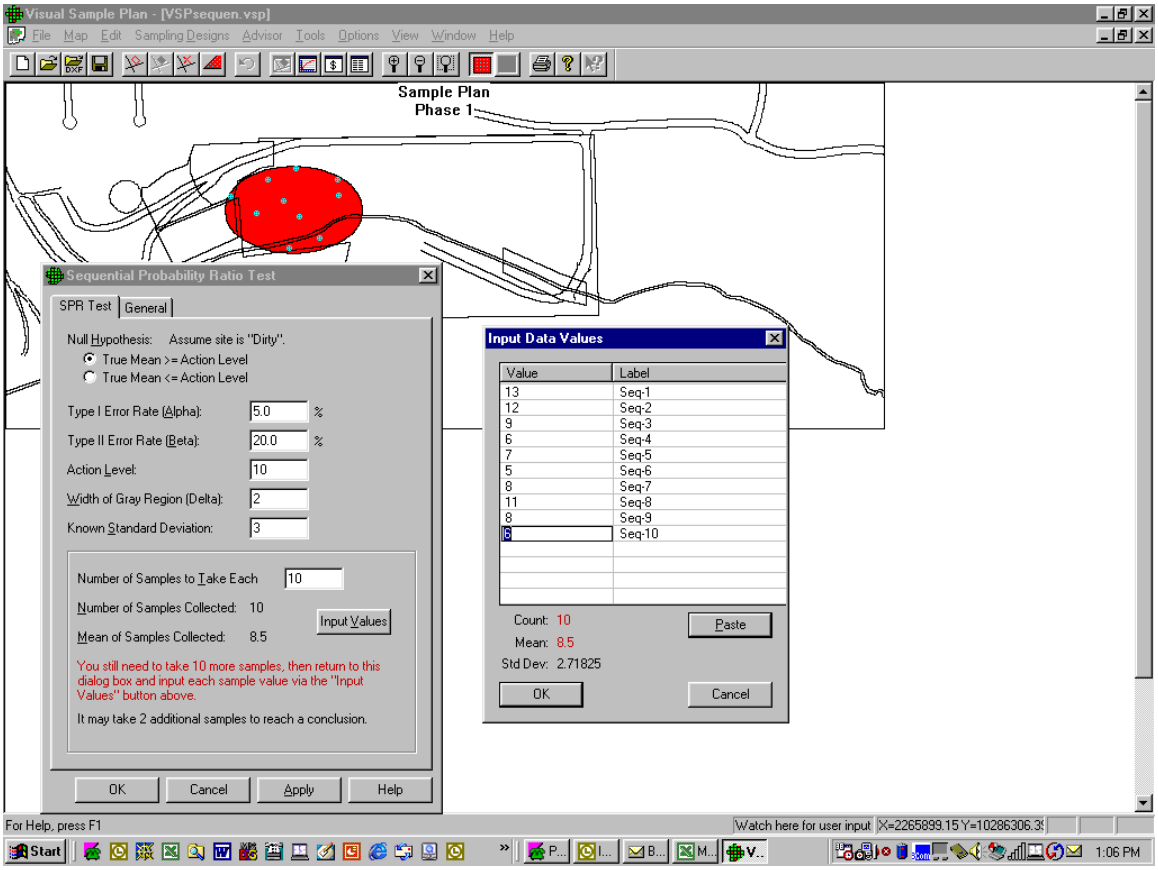


Figure 2. VSP Dialog Box

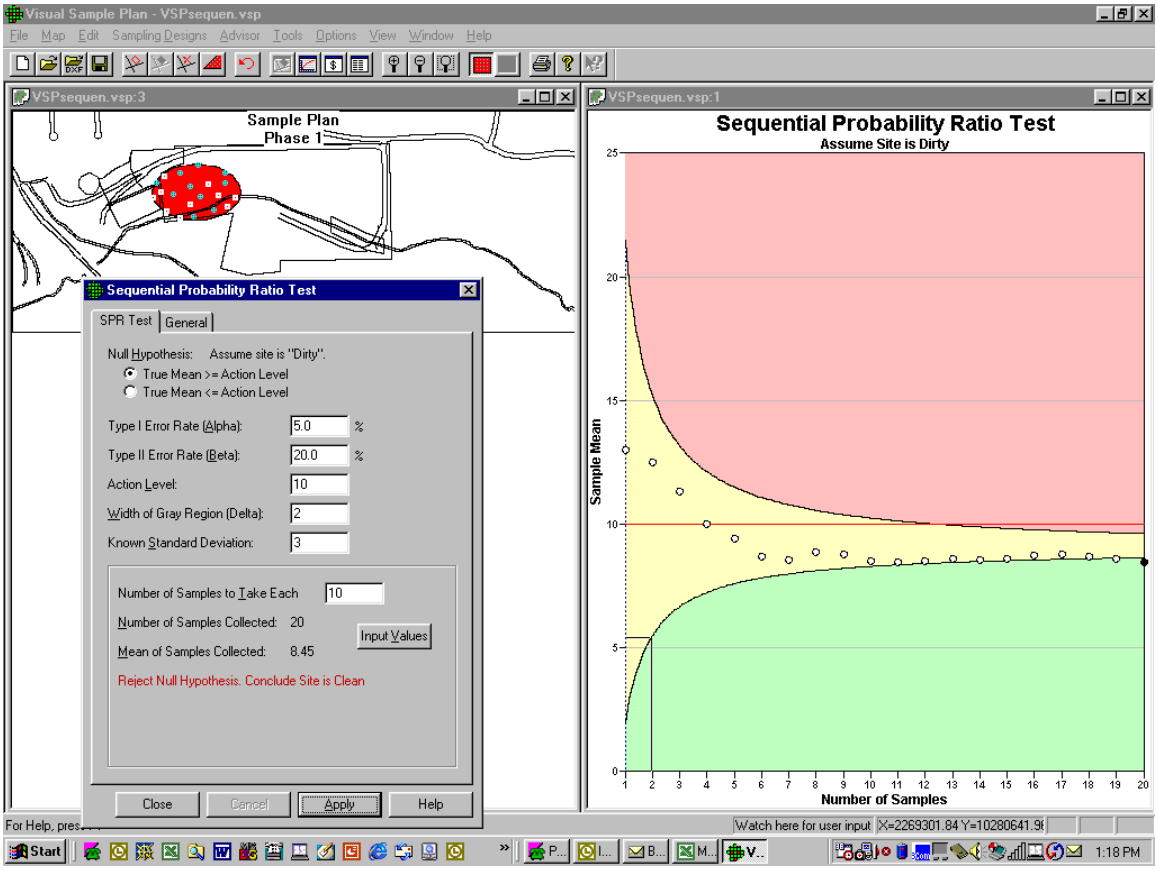


Figure 3. VSP Output

VSP Adaptive Sampling Module

EPA Office of Environmental Information (John Warren) is supporting the development of an Adaptive Cluster Sampling VSP module as part of an effort to incorporate all sampling designs listed in the EPA G5-S guidance [6,7]. Although this is still a work in progress, a brief summary of the method is presented. Currently, the VSP Adaptive Cluster Sampling module is only applicable when the sampling goal is to obtain an unbiased estimate of the contaminant mean concentration while identifying the boundaries of contaminated zones. The sampling area is first divided into a gridded set of sampling units. Then a random sample of all possible sampling units is selected to ensure that a confidence interval about the resulting mean will be a pre-specified width given several assumptions. A sample is obtained from each selected sampling unit and analyzed using in-field techniques. The resulting concentration data are input into VSP. VSP then identifies any sample results that exceed the decision threshold and selects adjacent sample units for the next round of sampling and analysis. This continues until either all sampling units adjacent to contaminated sampling units have been sampled or it becomes infeasible or not cost effective to continue.

As shown in Figure 4, VSP requires the following input for the Adaptive Cluster Sampling module.

- Whether a one-sided or two-sided confidence interval is of interest.
- Required confidence level.
- Maximum acceptable halfwidth of confidence interval.

- Estimated sampling and analytical standard deviation.
- Desired grid size for defining sampling units.
- Upper limit (threshold) for triggering adjacent sample unit sample collection.
- Whether 4 or 8 adjacent sampling units to be sampled if threshold exceeded.
- Sampling and analytical costs.
- Sample unit contaminant concentration results if unit is sampled.

Given this iterative input, VSP determines the number of samples and which sample units should be sampled for the initial round of sampling. A non-statistical comparison of the sample unit contaminant concentration against the triggering threshold is performed and adjacent sample units are recommended for the next round of sampling if the contaminant concentration exceeds the threshold. Although not currently available in VSP, eventually the unbiased estimate of the entire sample area mean will be calculated and a report will be automatically generated documenting the entire statistical sampling approach, all statistical formulas used, and the resulting estimates.

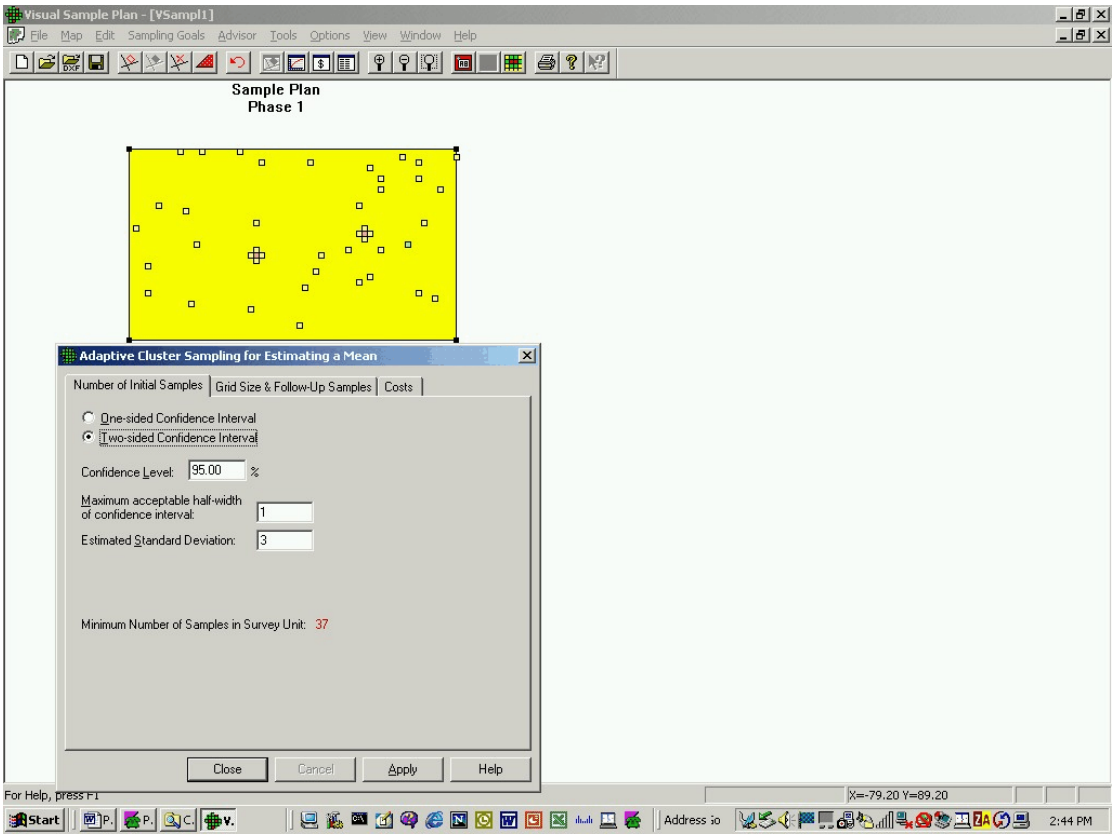


Figure 4. VSP Adaptive Clustering Input

Summary

Sequential and adaptive statistical sampling approaches appear to be useful when considering a Dynamic Work Plan. Visual Sample Plan modules for sequential and adaptive sampling are in the final stages of development. These sequential and adaptive modules are consistent with guidance on the DQO process and Dynamic Field Activities. Because VSP has been and continues to be developed using government funding, it is free to any government agency or contractor working on a US government site. The most recent VSP releases are found on http://dqo.pnl.gov/vsp and the updated version that will have sequential and adaptive sampling features is available upon request but will formally be released this summer.

References

[1] US EPA. 2002. *Guidance for Dynamic Field Activities*, Draft. Office of Emergency and Remedial Response, US Environmental Protection Agency, Washington, D.C. http://www.epa.gov/superfund/programs/dfa/guidoc.htm.

[2] US EPA. 2000a. *Guidance for the Data Quality Objectives Process - QA/G-4*. EPA/600/R-96/055, Office of Environmental Information, US Environmental Protection Agency, Washington, D.C.

[3] Davidson, JR, NL Hassig, JE Wilson, RO Gilbert. 2001. *Visual Sample Plan Version 1.0 User's Guide*. PNNL-13490, Pacific Northwest National Laboratory, Richland, Washington.

[4] Wald, Abraham. 1947. *Sequential Analysis*. J. Wiley & Sons, Inc., New York.

[5] Barnards, Rushton S. 1950. "On a Sequential t-Test." *Biometrika* 37:326-333.

[6] US EPA. 2000. *Guidance for Choosing a Sampling Design for Environmental Data Collection, QA/G-5S, Peer Review Draft.* Office of Environmental Information, US Environmental Protection Agency, Washington, D.C.

[7] Thompson, Steven K and George AF Seber. 1996. *Adaptive Sampling*. John Wiley & Sons, Inc., New York.